

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2004, 16:06:28 ; Search time 197 Seconds
(without alignments)

461.468 Million cell updates/sec

Title: US-10-015-085-4_COPY_1_158

Perfect score: 821

Sequence: 1 MKRVITLFAVLWGWSVNAW.....VAIKAGSLIAVLILQTNYY 158

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	821	100.0	166	2	Q9F5R9
2	821	100.0	300	1	FIMH_ECOLI
3	821	100.0	300	2	Q6JKU6
4	821	100.0	300	2	Q84DW2
5	821	100.0	300	2	Q9F5S0
6	817	99.5	300	2	Q7BYN8
7	817	99.5	300	2	Q83TY4
8	817	99.5	300	2	Q84DW3
9	817	99.5	300	2	Q83P73
10	815	99.3	300	2	Q9S494
11	815	99.3	300	2	Q9S496
12	815	99.3	300	2	Q9S497
13	815	99.3	300	2	Q9R5Y2
14	814	99.1	300	2	Q84DW4
15	814	99.1	300	2	Q9F5S1
16	813	99.0	300	2	Q6JKU4
17	812	98.9	300	2	Q6JKU5
18	812	98.9	300	2	Q6JKU3
19	812	98.9	300	2	Q6JKV4
20	811	98.8	300	2	Q6JKV7
21	811	98.8	300	2	Q7A8M4
22	811	98.8	300	2	Q8X8A6
23	810	98.7	300	2	Q9S495
24	809	98.5	300	2	Q6JKV8
25	809	98.5	300	2	Q84DW0
26	809	98.5	300	2	Q9F6Z7
27	809	98.5	303	2	Q8CVF8
28	808	98.4	300	2	Q84DW1
29	806	98.2	300	2	Q87634
30	806	98.2	303	2	Q9A505
31	804	97.9	300	2	Q84DW5

32	803	97.8	300	2	Q9S6R1	Q9S6R1	escherichia
33	798	97.2	300	2	Q6JKU8	Q6JKU8	escherichia
34	797	97.1	300	2	Q6JKU7	Q6JKU7	escherichia
35	714.5	87.0	243	2	Q08858	Q08858	klebsiella
36	501	61.0	129	2	Q9XBV8	Q9XBV8	escherichia
37	309.5	37.7	304	1	YDEQ_ECOLI	YDEQ_ECOLI	escherichia
38	309.5	37.7	304	2	Q7A5C8	Q7A5C8	escherichia
39	309.5	37.7	304	2	Q8XAX2	Q8XAX2	escherichia
40	307.5	37.5	304	2	Q8CW27	Q8CW27	escherichia
41	195.5	23.8	245	2	Q83RF2	Q83RF2	shigella fl
42	128	15.6	299	2	Q53298	Q53298	escherichia
43	127	15.5	299	1	SFAH_ECOL6	SFAH_ECOL6	escherichia
44	119	14.5	336	2	Q8CVB9	Q8CVB9	escherichia
45	117	14.3	299	2	Q46686	Q46686	escherichia

ALIGNMENTS

RESULT 1

Q9F5R9 PRELIMINARY; PRT; 166 AA.
 AC Q9F5R9;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE FimH.
 GN Name=fimH;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia;
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RA Harris S.L., Spears P.A., Havell E.A., Hamrick T.S., Horton J.R.,
 RA Orndorff P.E.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF306537; AAG30927.1; -
 DR HSSP; P08191; 10UN.
 DR InterPro; IPR008966; Adhes_bact.
 SQ SEQUENCE 166 AA; 17826 MW; 2E64EEA80AE286DC CRC64;

Query Match 100.0%; Score 821; DB 2; Length 166;
 Best Local Similarity 100.0%; Pred. No. 6.1e-67;
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLWGWSVNAWSPACKTANGTAIPGGGSANVYNLAPVNVGNLVVDLS 60
 Db 1 MKRVITLFAVLWGWSVNAWSPACKTANGTAIPGGGSANVYNLAPVNVGNLVVDLS 60

QY 61 TQIFCHNDYPTIIDYVTLQGSAYGGVLSNFSGTGVKSGSSYPPTTSETPRVYNSRT 120
 Db 61 TQIFCHNDYPTIIDYVTLQGSAYGGVLSNFSGTGVKSGSSYPPTTSETPRVYNSRT 120

QY 121 DKPWFVALYLPVSSAGGVAIKAGSLIAVLILQTNYY 158
 Db 121 DKPWFVALYLPVSSAGGVAIKAGSLIAVLILQTNYY 158

RESULT 2

FIMH_ECOLI
 ID FIMH_ECOLI STANDARD; PRT; 300 AA.
 AC P08191;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE FimH protein precursor.
 GN Name=fimH; OrderedLocusNames=b4320;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;

RN SEQUENCE FROM N.A.
 RP MEDLINE=88038337; PubMed=2890081;
 RA Klemm P., Christiansen G.;
 RT "Three fim genes required for the regulation of length and mediation
 RL of adhesion of *Escherichia coli* type 1 fimbriae.";
 RM Mol. Gen. Genet. 208:439-445(1987).
 [2] .
 RN SEQUENCE FROM N.A.
 RP STRAIN=K12 / MG1655;
 RC MEDLINE=95334362; PubMed=7610040;
 RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
 RA Blattner F.R.;
 RT "Analysis of the *Escherichia coli* genome VI: DNA sequence of the
 RL region from 92.8 through 100 minutes.";
 RM Nucleic Acids Res. 23:2105-2119(1995).
 [3] .
 RN CHARACTERIZATION.
 RP MEDLINE=90256291; PubMed=1971261;
 RC Krogfelt K.A., Bergmans H., Klemm P.;
 RA "Direct evidence that the FimH protein is the mannose-specific adhesin
 RL of *Escherichia coli* type 1 fimbriae.";
 RM Infect. Immun. 58:1995-1998(1990).
 CC -1- FUNCTION: Involved in regulation of length and mediation of
 CC adhesion of type 1 fimbriae (but not necessary for the production
 CC of fimbriae). Adhesin responsible for the binding to D-mannose. It
 CC is laterally positioned at intervals in the structure of the type
 CC 1 fimbriae. In order to integrate fimH in the fimbriae fimF and
 CC fimG are needed.
 CC -1- SIMILARITY: Belongs to the fimbrial protein family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; X05672; CAA29156.1; .
 CC EMBL; U14003; AAB97216.1; .
 CC EMBL; AE000502; AAC77276.1; .
 CC F01; S56545; S56545.
 CC PDB; 1KIU; X-ray; B/D/F/H/J/L/N/P=22-300.
 CC PDB; 1KLF; X-ray; B/D/F/H/J/L/N/P=22-300.
 CC PDB; 1QUN; X-ray; B/D/F/H/J/L/N/P=22-300.
 CC ECHOBASE; EB0311; .
 CC EcoGene; EG10315; fimH.
 CC InterPro; IPR008966; Adhes_bact.
 CC InterPro; IPR00259; Fimbrial.
 CC Pfam; PF00419; Fimbrial; 1.
 KW 3D-structure; Complete proteome; Fimbria; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 300 Potential.
 FT FAIN FimH protein.
 FT CONFLICT 197 197 P -> R (in Ref. 1).
 FT CONFLICT 222 222 T -> H (in Ref. 1).
 FT STRAND 23 26
 FT TURN 27 28
 FT STRAND 31 32
 FT TURN 34 35
 FT STRAND 38 43
 FT STRAND 48 48
 FT TURN 51 52
 FT STRAND 55 58
 FT HELIX 59 61
 FT TURN 62 62
 FT STRAND 63 66
 FT TURN 70 72
 FT STRAND 75 84
 FT HELIX 86 91
 FT STRAND 92 98
 FT TURN 99 100
 FT STRAND 101 105

FT STRAND 114 116
 FT STRAND 122 123
 FT STRAND 126 132
 FT TURN 134 135
 FT STRAND 139 141
 FT TURN 143 144
 FT STRAND 146 156
 FT STRAND 163 171
 FT STRAND 175 176
 FT STRAND 183 183
 FT STRAND 190 191
 FT STRAND 200 201
 FT STRAND 205 207
 FT STRAND 212 219
 FT STRAND 222 222
 FT TURN 225 226
 FT STRAND 229 231
 FT STRAND 238 238
 FT STRAND 240 240
 FT TURN 241 241
 FT STRAND 242 248
 FT TURN 249 250
 FT STRAND 251 253
 FT TURN 255 256
 FT STRAND 259 264
 FT STRAND 269 270
 FT STRAND 273 279
 FT STRAND 285 285
 FT STRAND 288 288
 FT STRAND 291 292
 FT STRAND 294 300
 SQ SEQUENCE 300 AA; 31473 MW; 939204A51658747D CRC64;
 Query Match 100.0%; Score 821; DB 1; Length 300;
 Best Local Similarity 100.0%; Pred. No. 1.2e-66;
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKRVITLFAVLMLGWSVNAWSPACKTANGTAIPIGGGSANVYNLAPVYVNGQNLVVDLS 60
 DB 1 MKRVITLFAVLMLGWSVNAWSPACKTANGTAIPIGGGSANVYNLAPVYVNGQNLVVDLS 60
 QY 61 TQIFCHNDYPTITDYVTLQSGAYGGVLSNFGTVKSGSSYPPTTSETPRVYVNSRT 120
 DB 61 TQIFCHNDYPTITDYVTLQSGAYGGVLSNFGTVKSGSSYPPTTSETPRVYVNSRT 120
 QY 121 DKFWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158
 DB 121 DKFWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158
 RESULT 3
 Q6JKV6 PRELIMINARY; PRT; 300 AA.
 ID Q6JKV6
 AC Q6JKV6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE FimH.
 GN Name=fimH;
 OS *Escherichia coli*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Escherichia*.
 ON NCBI_TaxID=562;
 RX SEQUENCE FROM N.A.
 RA Vandemaele F.J.; Hensen S.M., Goddeeris B.M.;
 RT "Conservation of deduced amino acid sequence of FimH among *Escherichia*
 RL coll. of bovine, porcine and avian disease origin.";
 RM Vet. Microbiol. 101:147-152(2004).
 RL EMBL; AY392512; AAR27559.1; .
 DR InterPro; IPR008966; Adhes_bact.
 DR InterPro; IPR00259; Fimbrial.

```
DR Pfam: PF00419; Fimbrial; 1.
SQ SEQUENCE 300 AA; 31457 MW; DC8CEBA50022C9BB CRC64;

Query Match      100.0%; Score 821; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.2e-66;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS 60
Db 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS 60

QY 61 TQIFCHNDYPETITDYVTLQRGSAYGVLNFSFGTVKYGSSYPFPPTTSETPRVYNSRT 120
Db 61 TQIFCHNDYPETITDYVTLQRGSAYGVLNFSFGTVKYGSSYPFPPTTSETPRVYNSRT 120

QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158
Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158

-RESULT 4
Q84DW2 PRELIMINARY; PRT; 300 AA.
AC Q84DW2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE FimH.
GN Name=fimH;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22545382; PubMed=12657207;
RA Vandemaële F., Vandekerckhove D., Vereecken M., Derijcke J.,
RA Dho-Moulin M., Goddeeris B.M.;
RT "Sequence analysis demonstrates the conservation of fimH and
RT variability of fimA throughout avian pathogenic Escherichia coli
RT (APEC).";
RL Vet. Res. 34:153-163 (2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Vandemaële F.J., Goddeeris B.M.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF490857; AA084616.1; -
DR HSSP; P08191; 1QUN
DR GO; GO:0002289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR008966; Adhes_bact.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
SQ SEQUENCE 300 AA; 31489 MW; 939219121PBF602D CRC64;

Query Match      100.0%; Score 821; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.2e-66;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS 60
Db 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS 60

QY 61 TQIFCHNDYPETITDYVTLQRGSAYGVLNFSFGTVKYGSSYPFPPTTSETPRVYNSRT 120
Db 61 TQIFCHNDYPETITDYVTLQRGSAYGVLNFSFGTVKYGSSYPFPPTTSETPRVYNSRT 120

QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158
Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158

RESULT 5
```

```
Q9F5S0 PRELIMINARY; PRT; 300 AA.
AC Q9F5S0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE FimH.
GN Name=fimH;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12;
RX MEDLINE=21289098; PubMed=11395476;
RA Harris S.L., Spears P.A., Havell E.A., Hamrick T.S., Horton J.R.,
RA Orndorff P.E.;
RT "Characterization of Escherichia coli type 1 pilus mutants with
RT altered binding specificities.";
RL J. Bacteriol. 183:4099-4102 (2001).
DR EMBL; AF306536; AAG30926.1; -
DR HSSP; P08191; 1QUN.
DR GO; GO:0002289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR008966; Adhes_bact.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
SQ SEQUENCE 300 AA; 31413 MW; 922904051758746D CRC64;

Query Match      100.0%; Score 821; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.2e-66;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS 60
Db 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS 60

QY 61 TQIFCHNDYPETITDYVTLQRGSAYGVLNFSFGTVKYGSSYPFPPTTSETPRVYNSRT 120
Db 61 TQIFCHNDYPETITDYVTLQRGSAYGVLNFSFGTVKYGSSYPFPPTTSETPRVYNSRT 120

QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158
Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNYY 158

RESULT 6
Q7BYN8 PRELIMINARY; PRT; 300 AA.
AC Q7BYN8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Minor fimbrial subunit, D-mannose specific adhesin.
GN Name=fimH; OrderedLocusNames=S4456;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=2457T;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786 (2003).
DR EMBL; AE016992; AAP19406.1; -
DR InterPro; IPR008966; Adhes_bact.
```

```

DR InterPro: IPR000259; Fimbrial.
DR Pfam: PF00419; Fimbrial; 1.
SQ SEQUENCE 300 AA; 31445 MW; E0278E990BC74884 CRC64;

Query Match
Best Local Similarity 99.5%; Score 817; DB 2; Length 300;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLGMGWSVNAWSPACKTANGTAIPIGGGSANVYNLAPVNVGNLWVDLS 60
Db 1 MKRAITLFAVLGMGWSVNAWSPACKTANGTAIPIGGGSANVYNLAPVNVGNLWVDLS 60
QY 61 TOIFCHNDYPETITDVTTLQSGAYGGVLSNFSGTVKYGSSYPPTTSETPRVYNSRT 120
Db 61 TOIFCHNDYPETITDVTTLQSGAYGGVLSNFSGTVKYGSSYPPTTSETPRVYNSRT 120
QY -121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158
Db 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158

RESULT 7.
Q83TY4 PRELIMINARY; PRT; 300 AA.
AC Q83TY4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE FimH.
GN Name=fimH;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22545392; PubMed=12657207;
RA Vandemaële F., Vanderkerchove D., Vereecken M., Derijcke J.,
RA Dho-Moulin M., Goddeeris B.M.;
RT "Sequence analysis demonstrates the conservation of fimH and
RT variability of fimA throughout avian pathogenic Escherichia coli
RT (APEC).";
RT Vet. Res. 34:153-163(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Vandemaële F.J., Goddeeris B.M.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=15172698;
RA Vandemaële F.J., Hensen S.M., Goddeeris B.M.;
RT "Conservation of deduced amino acid sequence of FimH among Escherichia
RT coli of bovine, porcine and avian disease origin.";
RL Vet. Microbiol. 101:147-152(2004).
DR EMBL; AF490845; AAO84604.1; -
DR EMBL; AF490847; AAO84606.1; -
DR EMBL; AF490848; AAO84607.1; -
DR EMBL; AF490850; AAO84609.1; -
DR EMBL; AF490851; AAO84610.1; -
DR EMBL; AF490852; AAO84611.1; -
DR EMBL; AF490853; AAO84612.1; -
DR EMBL; AF490855; AAO84614.1; -
DR EMBL; AF490858; AAO84617.1; -
DR EMBL; AF490859; AAO84618.1; -
DR EMBL; AF490860; AAO84619.1; -
DR EMBL; AF490861; AAO84620.1; -
DR EMBL; AF490862; AAO84621.1; -
DR EMBL; AF490863; AAO84622.1; -
DR EMBL; AF490864; AAO84623.1; -
DR EMBL; AY392515; AAR27562.1; -
DR HSP; P08191; 1QUN.
DR GO; GO:000289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.

DR InterPro: IPR008966; Adhes.bact.
DR InterPro: IPR000259; Fimbrial.
DR Pfam: PF00419; Fimbrial; 1.
SQ SEQUENCE 300 AA; 31445 MW; 9E0904A513D3747D CRC64;

Query Match
Best Local Similarity 99.5%; Score 817; DB 2; Length 300;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLGMGWSVNAWSPACKTANGTAIPIGGGSANVYNLAPVNVGNLWVDLS 60
Db 1 MKRAITLFAVLGMGWSVNAWSPACKTANGTAIPIGGGSANVYNLAPVNVGNLWVDLS 60
QY 61 TOIFCHNDYPETITDVTTLQSGAYGGVLSNFSGTVKYGSSYPPTTSETPRVYNSRT 120
Db 61 TOIFCHNDYPETITDVTTLQSGAYGGVLSNFSGTVKYGSSYPPTTSETPRVYNSRT 120
QY 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158
Db 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158

RESULT 8.
Q84DW3 PRELIMINARY; PRT; 300 AA.
AC Q84DW3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE FimH.
GN Name=fimH;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22545382; PubMed=12657207;
RA Vandemaële F., Vanderkerchove D., Vereecken M., Derijcke J.,
RA Dho-Moulin M., Goddeeris B.M.;
RT "Sequence analysis demonstrates the conservation of fimH and
RT variability of fimA throughout avian pathogenic Escherichia coli
RT (APEC).";
RT Vet. Res. 34:153-163(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Vandemaële F.J., Goddeeris B.M.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF490856; AAO84615.1; -
DR HSP; P08191; 1QUN.
DR GO; GO:000289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR008966; Adhes.bact.
DR InterPro: IPR000259; Fimbrial.
DR Pfam: PF00419; Fimbrial; 1.
SQ SEQUENCE 300 AA; 31426 MW; 8E02C3A8A50819DA CRC64;

Query Match
Best Local Similarity 99.4%; Score 817; DB 2; Length 300;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLGMGWSVNAWSPACKTANGTAIPIGGGSANVYNLAPVNVGNLWVDLS 60
Db 1 MKRAITLFAVLGMGWSVNAWSPACKTANGTAIPIGGGSANVYNLAPVNVGNLWVDLS 60
QY 61 TOIFCHNDYPETITDVTTLQSGAYGGVLSNFSGTVKYGSSYPPTTSETPRVYNSRT 120
Db 61 TOIFCHNDYPETITDVTTLQSGAYGGVLSNFSGTVKYGSSYPPTTSETPRVYNSRT 120
QY 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158
Db 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158

```

```

RESULT 9
Q83P73 ID Q83P73 PRELIMINARY; PRT; 300 AA.
AC Q83P73
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Minor fimbrial subunit, D-mannose specific adhesin.
GN Name=fim5; OrderedLocusNames=SF4200;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12394590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL; AE015428; AAN45621.1; -.
DR HSP; P08191; IQUN.
DR GO; GO:0009289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR008966; Adhes.bact.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
KW Complete proteome.
SQ SEQUENCE 300 AA; 31445 MW; B027E890BC74884 CRC64;

Query Match 99.5%; Score 817; DB 2; Length 300;
Best Local Similarity 99.4%; Pred. No. 2.8e-66;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
DB 1 MKRAITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
QY 61 TOIFCHNDYPETITDYVTLQRGSAVGVLNFSGTGVKSGSYPPFTTSETPRVYNSRT 120
DB 61 TOIFCHNDYPETITDYVTLQRGSAVGVLNFSGTGVKSGSYPPFTTSETPRVYNSRT 120
QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
DB 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 10
Q8S494 ID Q8S494 PRELIMINARY; PRT; 300 AA.
AC Q8S494
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE FimH.
GN Name=fimH;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=20327582; PubMed=10869080;
RA Hamrick T.S., Harris S.L., Spears P.A., Havell E.A., Horton J.R.,
RA Russell P.W., Orndorff P.E.;
RT "Genetic characterization of Escherichia coli type 1 pilus adhesin
RT mutants and identification of a novel binding phenotype."
RL J. Bacteriol. 182:4012-4021(2000).
DR EMBL; AF154926; AAD44320.1; -.
DR HSP; P08191; IQUN.
DR GO; GO:0009289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR008966; Adhes.bact.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
KW Complete proteome.
SQ SEQUENCE 300 AA; 31499 MW; 3ADF97A85F8CE478 CRC64;

Query Match 99.3%; Score 815; DB 2; Length 300;
Best Local Similarity 99.4%; Pred. No. 4.2e-66;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
DB 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
QY 61 TOIFCHNDYPETITDYVTLQRGSAVGVLNFSGTGVKSGSYPPFTTSETPRVYNSRT 120
DB 61 TOIFCHNDYPETITDYVTLQRGSAVGVLNFSGTGVKSGSYPPFTTSETPRVYNSRT 120
QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
DB 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158

```

```

RT mutants and identification of a novel binding phenotype.";
RL J. Bacteriol. 182:4012-4021(2000).
DR EMBL; AF154928; AAD44322.1; -.
DR HSP; P08191; IQUN.
DR GO; GO:0009289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR008966; Adhes.bact.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
SQ SEQUENCE 300 AA; 31516 MW; B6073DF68366886E CRC64;

Query Match 99.3%; Score 815; DB 2; Length 300;
Best Local Similarity 99.4%; Pred. No. 4.2e-66;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
DB 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
QY 61 TOIFCHNDYPETITDYVTLQRGSAVGVLNFSGTGVKSGSYPPFTTSETPRVYNSRT 120
DB 61 TOIFCHNDYPETITDYVTLQRGSAVGVLNFSGTGVKSGSYPPFTTSETPRVYNSRT 120
QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
DB 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 11
Q8S496 ID Q8S496 PRELIMINARY; PRT; 300 AA.
AC Q8S496
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE FimH.
GN Name=fimH;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=20327582; PubMed=10869080;
RA Hamrick T.S., Harris S.L., Spears P.A., Havell E.A., Horton J.R.,
RA Russell P.W., Orndorff P.E.;
RT "Genetic characterization of Escherichia coli type 1 pilus adhesin
RT mutants and identification of a novel binding phenotype."
RL J. Bacteriol. 182:4012-4021(2000).
DR EMBL; AF154926; AAD44320.1; -.
DR HSP; P08191; IQUN.
DR GO; GO:0009289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR008966; Adhes.bact.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
KW Complete proteome.
SQ SEQUENCE 300 AA; 31499 MW; 3ADF97A85F8CE478 CRC64;

Query Match 99.3%; Score 815; DB 2; Length 300;
Best Local Similarity 99.4%; Pred. No. 4.2e-66;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
DB 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
QY 61 TOIFCHNDYPETITDYVTLQRGSAVGVLNFSGTGVKSGSYPPFTTSETPRVYNSRT 120
DB 61 TOIFCHNDYPETITDYVTLQRGSAVGVLNFSGTGVKSGSYPPFTTSETPRVYNSRT 120
QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
DB 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158

```

Db 121 DKPWPVALYLPVSAGGVAIKAGSLIAVLLRQTNNY 158

RESULT 12

Q9S497 PRELIMINARY; PRT; 300 AA.

ID Q9S497

AC Q9S497

DT 01-WAY-2000 (TrEMBLrel. 13, Created)

DT 01-WAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE FimH.

DE Name=fimH;

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

OX [1]

RP SEQUENCE FROM N.A.

RP STRAIN=K12;

RX MEDLINE=20327582; PubMed=10869080;

RA Hamrick T.S., Harris S.L., Spears P.A., Havell E.A., Horton J.R.;

RA Russell P.W., Orndorff P.E.;

RT "Genetic characterization of *Escherichia coli* type 1 pilus adhesin mutants and identification of a novel binding phenotype.";

RL J. Bacteriol. 182:4012-4031(2000).

RL EMBL; AF154925; AAD44319.1; -

DR HSP; P08191; 1QUN.

DR GO; GO:0009289; C:fimbria; IEA.

DR GO; GO:0007155; P:cell adhesion; IEA.

DR InterPro; IPR008966; Adhes.bact.

DR InterPro; IPR00259; Fimbrial.

DR Pfam; PF00419; Fimbrial; 1.

DR SQ SEQUENCE 300 AA; 31517 MW; D29340A1575A2358 CRC64;

Query Match 99.3%; Score 815; DB 2; Length 300;

Best Local Similarity 99.4%; Pred.No. 4.2e-66;

Matches 157; Conservative 0; Mismatches 1; Indels 0;

QY 1 MKRVITLFAVLLMGWSVNNASFPACKTANGTAIPIGGSANVYVNLAPVWVGQNLV

Db 1 MKRVITLFAVLLMGWSVNDNSFPACKTANGTAIPIGGSANVYVNLAPVWVGQNLV

QY 61 TQIFCHNDYPETITDYTLQSGSAYGVGLNFSGTGVKYGSSYPPTTSETPRVVY

Db 61 TQIFCHNDYPETITDYTLQSGSAYGVGLNFSGTGVKYGSSYPPTTSETPRVVY

QY 121 DKPWPVALYLPVSAGGVAIKAGSLIAVLLRQTNNY 158

Db 121 DKPWPVALYLPVSAGGVAIKAGSLIAVLLRQTNNY 158

RESULT 13

Q9S5Y2 PRELIMINARY; PRT; 300 AA.

ID Q9S5Y2

AC Q9S5Y2

DT 01-WAY-2000 (TrEMBLrel. 13, Created)

DT 01-WAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE FIMH SUBUNIT=MANNOS- sensitive type 1 fimbrial adhesin.

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

OX [1]

RP SEQUENCE FROM N.A.

RP MEDLINE=94131954; PubMed=7905476;

RA Sokurenko B.V., Courtney H.S., Ohman D.E., Klemm P., Hasty D.L.;

RL J. Bacteriol. 176:748-755(1994).

DR HSP; P08191; 1QUN.

DR GO; GO:0009289; C:fimbria; IEA.

DR GO; GO:0007155; P:cell adhesion; IEA.

DR InterPro; IPR008966; Adhes.bact.

DR InterPro; IPR00259; Fimbrial.

```
Q9F5S1
ID Q9F5S1 PRELIMINARY; PRT; 300 AA.
AC Q9F5S1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE FimH.
DE GN
DE GN Names=fimH;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=21289098; PubMed=11395476;
RA Harris S.L., Spears P.A., Havell E.A., Hamrick T.S., Horton J.R.,
RA Orndorff P.E.;
RT "Characterization of Escherichia coli type 1 pilus mutants with
RT altered binding specificities.";
RL J. Bacteriol. 183:4099-4102(2001).
DR EMBL; AF306535; AAC30925.1; -
DR HSP; P08191; IQUN
DR GO; GO:0009289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR008966; Adhes_bact.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
SQ SEQUENCE 300 AA; 31457 MW; 9F7B0A35FF51F938 CRC64;
```

Query Match 99.1%; Score 814; DB 2; Length 300;
Best Local Similarity 99.4%; Pred. No. 5.2e-66;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
OY 1 MKRVTILFAVLLMGVSNVWSPACKTANGTAIPIGGSSANVYNLAPVNVGNLVVDLS 60
DB 1 MKRVTILFAVLLMGVSNVWSPACKTANGTAIPIGGSSANVYNLAPVNVGNLVVDLS 60

OY 61 TQIFCHNDYPETITDYVTLPQGSAYGGVLSNFSGTGVKSGSSYPPTTSETPRVYNSRT 120
DB 61 TQIFCHNDYPETITDYVTLPQGSAYGGVLSNFSGTGVKSGSSYPPTTSETPRVYNSRT 120

OY 121 DKRPVALYITPVSSAGGVAIKAGSLIAVLILRQTNNY 158
DB 121 DKRPVALYITPVSSAGGVAIKAGSLIAVLILRQTNNY 158
```

Search completed: November 27, 2004, 16:10:47
Job time : 204 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2004, 16:06:28 ; Search time 39 Seconds

(without alignments)
389.801 Million cell updates/sec

Title: US-10-015-085-4_COPY_1_158

Perfect score: 821

Sequence: 1 MKRVITLFAVLLMGWSNAN.....VAIKAGSLIAVLLRQTNNY 158

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 9621673 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

*Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- PIR 78:**
1: pir1:**
2: pir2:**
3: pir3:**
4: pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	821	100.0	300	S56545	fimbrial protein f
2	811	98.8	300	G91017	hypothetical prote
3	811	98.8	300	B86130	hypothetical prote
4	714.5	87.0	302	A32801	fimbrial adhesin p
5	309.5	37.7	304	A64304	probable fimbrial
6	309.5	37.7	304	F85725	probable adhesin.
7	309.5	37.7	304	C90892	probable adhesin
8	128	15.6	299	C49233	S fimbrial adhesin
9	127	15.5	291	S15927	sfaH protein precu
10	117	14.3	299	I76900	F1652 minor fimbri
11	86.5	10.5	682	TQ0420	beta-1,3-glucanase
12	85	10.4	317	T39869	probable lysophosp
13	82.5	10.0	354	A82593	uroporphyrinogen d
14	82.5	10.0	408	T22801	hypothetical prote
15	80.5	9.8	473	A59200	acid phosphatase (
16	80.5	9.8	482	C84330	pantothenate perme
17	80.5	9.8	558	B99494	dihydroxy-acid deh
18	80	9.7	1114	JH0284	125K surface anti
19	79.5	9.7	432	T71683	NADH2 dehydrogenas
20	79.5	9.7	653	A60961	leishmanolysin (EC
21	79.5	9.7	824	I S50767	S-receptor kinase
22	78.5	9.6	343	H95879	probable sugar ABC
23	78.5	9.6	451	C2 F89932	hypothetical prote
24	78	9.5	184	C86320	hypothetical prote
25	78	9.5	282	AB1359	biogenesis of thyl
26	78	9.5	471	C64665	hypothetical prote
27	77.5	9.4	412	C75279	hypothetical prote
28	77.5	9.4	1337	C64993	hypothetical prote
29	77.5	9.4	1534	A85862	hypothetical prote

30	77.5	9.4	1534	2	G91017	probable membrane
31	77	9.4	309	2	AD0221	probable sugar ABC
32	77	9.4	471	2	B71849	probable outer mem
33	77	9.4	630	2	S37663	leptomeric polypro
34	77	9.4	1154	1	VGIHIB	E2 glycoprotein pr
35	76	9.3	246	2	C70527	hypothetical prote
36	75.5	9.2	378	2	S00842	leukosialin precu
37	75.5	9.2	383	2	G84248	iron-binding prote
38	75.5	9.2	485	2	T51094	acid phosphatase (
39	75	9.1	276	2	G86130	hypothetical prote
40	75	9.1	276	2	D91289	hypothetical prote
41	75	9.1	849	2	D82510	chitinase VCA0027
42	75	9.1	1041	2	S55862	probable membrane
43	74.5	9.1	341	2	E96019	probable sugar upt
44	74.5	9.1	352	2	A95098	hypothetical prote
45	74.5	9.1	396	2	AE2137	hypothetical prote

ALIGNMENTS

RESULT 1

S56545
fimbrial protein fimH precursor, type 1 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S56545; B65246; S09563; A36967
R:Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res 23, 2105-2119, 1995
A>Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92.
A:Reference number: S56314; MUID:95334362; PMID:7610040
A:Accession: S56545
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-300 <BUR>
A:Cross-references: UNIPROT:P08191; EMBL:U14003; NID:G1263172; PIDN:AAA97216.1; PID:G537
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B65246
A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-300 <BLAT>
A:Cross-references: GB:AB000502; GB:U00096; NID:G2367374; PIDN:AAC77276.1; PID:G1790775;
A:Experimental source: strain K-12, substrain MG1655
R:Klemm, P.; Christiansen, G.
Mol. Gen. Genet 208, 439-445, 1987
A>Title: Three fim genes required for the regulation of length and mediation of adhesion
A:Reference number: S07321; MUID:88038337; PMID:2890081
A:Accession: S09563
A:Molecule type: DNA
A:Residues: 1-196 'R', 198-221 'H', 223-300 <KLE>
A:Cross-references: EMBL:X05672; NID:G41463; PIDN:CAA29156.1; PID:G41466
R:Sokurenko, E.V.; Courtney, H.S.; Ohman, D.E.; Klemm, P.; Hasty, D.L.
J. Bacteriol. 176, 748-755, 1994
A>Title: FimH family of type 1 fimbrial adhesins: functional heterogeneity due to minor
A:Reference number: A36967; MUID:94131954; PMID:7905476
A:Accession: A36967
A>Status: not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-78 'R', 80-221 'H', 223-300 <SOK>
A>Note: sequence extracted from NCBI backbone (NCBIP:143314)
C:Genetics:
A:Gene: fimH
C:Function:
A:Description: involved in longitudinal regulation and mannose-specific adhesion
A>Note: not necessary for the production of fimbriae
A>Note: controls length and number of fimbriae
C:Superfamily: fimbrial protein fimH
C:Keywords: fimbria

```

RESULT 3
B86130      hypothetical protein fimH [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: B86130
RifPerna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B86130
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-300 <STO>

```

```

RESULT 5
A64904
C:Species: Escherichia coli
C:Strain: probable fimbrial protein b1502 - Escherichia coli (strain K-12)
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change
C:Accession: A64904
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Bu
.A.; Rose, D.J.; Mau, B.; Shaoy, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference Number: A64720; MUID:97426617; PMID:9278503
A:Accession: A64904
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-304 <BLAT>
A:Cross-references: UNIPROT: P77588; GR:AF000247; NID:

```

A:Residues: 1-304 <HAY>
A:Cross-references: UNIPROT:Q8XAX2; GB:BA000007; PIDN:BA835530.1; PID:GL3361573; GSPDB:B
A:Experimental source: strain O157:H7, substrain RMD 050952
C:Genetics:
A:Gene: ECs2107
C:Superfamily: fimbrial protein fimH

Query Match 37.7%; Score 309.5; DB 2; Length 304;
Best Local Similarity 44.1%; Pred. No. 5.7e-21;
Matches 63; Conservative 29; Mismatches 48; Indels 3; Gaps 2;

QY 7 LFAVLMLGMSVNAVSFACTANGTAIPICGGSANYVNLAPVVNVGQNLVVDLSTQIFCH 66
DB 10 LFGIYLLMAGKVFAFSCNVDDGSS--IGAGTTSVVVNLDPVIQPCQNLVVDLSHISCW 67
QY 67 NDYPERI-TDVTYLGSAAYGGVLSNFSTKVSYSSSYPPTTSETPRVVNSRTDKWP 125
DB 68 NDYGGWYTDHNLVQGSAPASLQSKYSLYNNVTFPPLTTNTNVLDIGDKTPMLP 127
QY 126 VALYLTVPSSAGGVAIKAGSLIA 148
DB 128 LKLYITPVGAAGGVWIKAGEVIA 150

RESULT 8
C49233
S fimbrial adhesin minor subunit SfaH - Escherichia coli
C:Species: Escherichia coli
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: C49233
R:Hacker, J.; Kestler, H.; Hoschutsky, H.; Jann, K.; Lottspeich, F.; Korhonen, T.K.
A:Title: Cloning and characterization of the S fimbrial adhesin II complex of an Escherichia coli strain O157:H7
A:Reference number: A49233; MUID:93138776; PMID:8093693
A:Contents: O18:K1
A:Accession: C49233
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-299 <HAC>
A:Cross-references: UNIPROT:Q83298; GB:S53210; NID:G2644033; PIDN:AAB25047.1; PID:G264036
A:Note: sequence extracted from NCBI backbone (NCBIN:123181, NCBIP:123186)
C:Superfamily: fimbrial protein fimH

Query Match 15.6%; Score 128; DB 2; Length 299;
Best Local Similarity 25.8%; Pred. No. 0.00027;
Matches 39; Conservative 24; Mismatches 78; Indels 10; Gaps 4;

QY 13 MGMSVNAVSFACTANGTAIPICGGSANYVNLAPVVNVGQNL-VVDLSTQIFCH-DYP 70
DB 1 MAYSQSPALLCN-NQTGOVFNSTGTSFRVNVSPVQYDKSISVLDSLQVSCQNEST 59
QY 71 ETITDVTYLGSAAYGGVLS-----NFGTVKYSYSSSYPPTTSETPRVVNSRTDKP 123
DB 60 GQNVYDLKILKSGFSPALDTKTYGRLDTSRPTGYARQLPQLQFDLQVTEAFYQYGVWKP 119
QY 124 WPVALYITPVSSAGGVAIKAGSLIAVLRLQ 154
DB 120 FPAKLYLPAPGVFGKVINNGDLATLYNKK 150

RESULT 9
S15927
sfaH protein precursor - Escherichia coli
C:Species: Escherichia coli
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C:Accession: S15927; S06195
R:Schmoll, T.; Hoschutsky, H.; Morschhauser, J.; Lottspeich, F.; Jann, K.; Hacker, J.
Mol. Microbiol. 3, 1735-1744, 1989
A:Title: Analysis of genes coding for the sialic acid-binding adhesin and two other minor adhesins of Escherichia coli O157:H7
A:Reference number: S15925; MUID:90158121; PMID:2576095
A:Accession: S15927
A:Molecule type: DNA
A:Residues: 1-291 <SCH>

A;Cross-references: UNIPROT:PI3431; EMBL:X16664; NID:g42955; PIDN:CAA34654.1; PID:g42962
C;Genetics:
A;Gene: sfah
C;Superfamily: fimbrial protein fimH
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-291/Product: sfah protein #status predicted <MAT>

Query Match 15.5%; Score 127; DB 2; Length 291;
Best Local Similarity 25.8%; Pred. No. 0.00032;
Matches 39; Conservative 24; Mismatches 78; Indels 10; Gaps 4;

QY 13 MGSVNVAMSFACKTANGTAIPIGGSANVYVNLAPVNVGNQL-VVDLSTQIFCHN-DYP 70
Db 1 MAYSQSFALLCRN-NQTQFNSGDTSPRVNVPVVEYDKSISVLDLSQLVSCQNEBST 59
QY 71 ETITDYVTLQGSAYGGVLS-----NFSGTVKYSGSSYPPTTSETPRVYNSRTDKP 123
Db 60 GQNDYLYKILKSGFSPALDRTYGRDLPTSRPTGVARQLPQLQDVTAEFYQYGVWKP 119
QY 124 WPVALYLPVSSAGGVAIKAGSLIAVLILRQ 154
Db 120 FPAKWLYPEPGVFGKVINNGDLATLYVVK 150

RESULT 10
I76900
F1852 minor fimbrial chain H - Escherichia coli
C;Species: Escherichia coli
C;Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 09-Jul-2004
C;Accession: I76900; I80334
R;Harel, J.; Jacques, M.; Fairbrother, J.M.; Bosse, M.; Forget, C.
Microbiology 141, 221-228, 1995
A;Title: Cloning of determinants encoding F185(2) fimbriae from porcine septicaemic Esch
A;Reference number: I5911; MUID:95202083; PMID:7894716
A;Accession: I76900
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-299 <RES>
A;Cross-references: UNIPROT:O46686; EMBL:U09804; NID:g967126; PIDN:AAA74946.1; PID:g9671
R;van Die, I.; Kramer, C.; Hacker, J.; Bergmans, H.; Jongen, W.; Hoekstra, W.
Res. Microbiol. 142, 653-658, 1991
A;Title: Nucleotide sequence of the genes coding for minor fimbrial subunits of the F1C
A;Reference number: I59446; MUID:92073661; PMID:1683712
A;Accession: I80334
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-299 <RE2>
A;Cross-references: GB:S68237; NID:g239708; PIDN:AAB20439.1; PID:g239711
C;Genetics:
A;Gene: foch
C;Superfamily: fimbrial protein fimH

Query Match 14.3%; Score 117; DB 2; Length 299;
Best Local Similarity 25.8%; Pred. No. 0.0028;
Matches 39; Conservative 27; Mismatches 75; Indels 10; Gaps 6;

QY 13 MGSVNVAMSFACKTANGTAIPIGGSANVYVNLAPVNVGNQL-VVDLSTQIFCHN-DYP 70
Db 1 MAYSQSFALLCRN-NQTQFNSGDTSPRVNVPVVEYDKSISVLDLSQLVSCQNEBST 59
QY 71 ETITDYVTLQGSAYGGVLS-----SNFSGTVKYSGSSYPPTTSET-PRVYV-NSRTDKP 123
Db 60 GQNDYLYRQGTGSPSLDAKTYGRDLPTSRPTGVARQLPQLQDVTAEFYQYGVWKP 119
QY 124 WPVALYLPVSSAGGVAIKAGSLIAVLILRQ 154
Db 120 FPAKWLYPEPGVFGKLIHAGSLVATVYVK 150

RESULT 11
JQ0420
beta-1,3-glucanase A1 precursor - Bacillus circulans
C;Species: Bacillus circulans

A;Cross-references: UNIPROT:P23903; GB:M34503; NID:g142972; PIDN:AAA22474.1; PID:g142973
A;Residues: 1-682 <YAH>
A;Cross-references: UNIPROT:P23903; GB:M34503; NID:g142972; PIDN:AAA22474.1; PID:g142973
A;Experimental source: strain WL-12
C;Comment: This enzyme, together with chitinase, is crucial for hydrolyzing yeast and fur
C;Genetics:
A;Gene: glcA
F;1-38/Domain: signal sequence #status predicted <SIG>
F;39-682/Product: beta-1,3-glucanase A1 #status predicted <MAT>

Query Match 10.5%; Score 86.5; DB 2; Length 682;
Best Local Similarity 22.2%; Pred. No. 4.7;
Matches 47; Conservative 30; Mismatches 56; Indels 79; Gaps 9;

QY 1 MKRVTILFAVLLMGVSNVAMSFACKTANGTAI-----PI----- 34
Db 12 MKKVLGLFLVWMLASVGLPTSKVQAAGTIVTSMYFSPADGPVTSKSGVKASYGVFM 71
QY 35 ---GGGSA---NVYVNLAPVNVGNQLV-VD-----LSTQIFCH-----NDYPETITDY 76
Db 72 PKFNGSATWVDYSDGVNVKVVGNVNDIDQAGVYVQNVNHWSDGGFNGYWTLSAT 131
QY 77 VTLQGRSAYGGV-----LSNFSGTVKYSGSSYPPTTSS 109
Db 132 TEIQLYSKANGVKLEQLVFQINQKTTITAMNPTQGPQTASFTG-----GAGFTYPTFN 186
QY 110 ETRPVVNSRTDYPWVALYLPVSSAGGVAI 141
Db 187 NDSAVTYEAVD---DLKRVYVKNVSSSWIDI 215

RESULT 12
T39869
Probable lysophospholipase (EC 3.1.1.5) SPBC1E8.05 [similarity] - fission yeast (Schizos
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T39869; T42012
R;Volckaert, G.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z21887
A;Accession: T39869
A;Status: translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-317 <VOL>
A;Cross-references: UNIPROT:O42970; EMBL:AL021746; NID:g2842520; PIDN:CAA16838.1; PID:g28
A;Experimental source: strain 972h; cosmid c1E8
R;Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A;Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
A;Reference number: Z17323; MUID:98162722; PMID:9501991
A;Accession: T42012
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 6-123, K' 125-317, 'IVVS' <YOS>
A;Cross-references: EMBL:D89103; NID:g1749413; PIDN:BAAL3766.1; PID:g1749414
A;Experimental source: strain PR745
C;Genetics:
A;Gene: SPDB:SPBC1E8.05
A;Map position: 2
C;Superfamily: yeast glucan 1,4-alpha-glucosidase homolog; glucan 1,4-alpha-glucosidase
C;Keywords: carboxylic ester hydrolase

Query Match 10.4%; Score 85; DB 2; Length 317;
Best Local Similarity 26.0%; Pred. No. 2.6;
Matches 38; Conservative 24; Mismatches 46; Indels 38; Gaps 7;

QY 9 AVLLMGWSNVNWSFACKTANGTA.PI--GGGSANVYVNLAPVNVVGNQL-----VVDLS 60
Db 27 AVLQEGGVNTVSNLSTVTTLTYRGNSA-----LTPETIASIDNTGTGLWNIA 81
QY 61 TQIFCHNDY-----PETITDVTTLQSGS-----AYGVLNFPN-----CTVK 97
Db 82 TYVEAADYLLGLSFDGGETYSQYFTLQASTCTISTSSLSYSGTSTSTSIASMTG--- 139
QY 98 YSGSSYFPPTTSEPRVYVNSRTDKP 123
Db 140 RTSSSYFTTSSSTPSSSSSSSSSSSP 165

RESULT 13
A82693
uoroporphyrinogen decarboxylase XF1332 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: A82693
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen-
tature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82615; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: A82693
A:Status: preliminary
A:Molecule type: DNA
A:Map position: 1
A:Residues: 1-354 <SLM>
A:Cross-references: UNIPROT:Q9PDP7; GB:AF003966; GB:AF003849; NID:99106327; PIDN:AAF8414
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A-
briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H-
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2003
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm-
u.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laig-
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E-
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Mitacca, E.C.; Miyaki, C.V.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A-
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak-
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silvair-
M.; Tshuko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z-
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1332
C:Superfamily: uoroporphyrinogen decarboxylase

Query Match 10.0%; Score 82.5; DB 2; Length 354;
Best Local Similarity 23.2%; Pred. No. 5;
Matches 44; Conservative 24; Mismatches 57; Indels 65; Gaps 10;

QY 9 AVLLMGWSNVNWSFACKTANGTA.PI--GGGSANVYVNL-----APVNVGNLWVDLSQ 62
Db 136 SVFLIFSGSFWTLAC-----YMEGGSKSEYAIKAMAFNAP-----DVLHQ 178

QY 63 IFCHNDYPETITDVTTLQSGS-----AYGVLNFPN-----AYGVLNFPN----- 107
Db 179 LL--NTVNAVSYLAQAQAQAQALQVFDTWGGLS-----PMYRTSLPYLTRIAREL 232

QY 108 -----TSEPRVYV-----SRTDX-----PWPVALYLPVSSAGGVAIKAGSLI 147
Db 233 ERGTGTERTPLVFGKNGBYIAELAMSGTEAVGVVDVTDIELEDAARRSNGVALQ--GNLD 291

QY 148 AVILRQTN 157
Db 292 PATLYGTENN 301

RESULT 14
T22801
hypotheical protein F56H6.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T22801
R:Kershaw, J.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19617
A:Accession: T22801
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-408 <WIL>
A:Cross-references: UNIPROT:O45586; EMBL:Z81553; PIDN:CAB04497.1; GSPDB:GN00019; CESP:FE
A:Experimental source: clone F56H6
C:Genetics:
A:Gene: CESP:F56H6.8
A:Map position: 1
A:Introns: 67/3; 159/1

Query Match 10.0%; Score 82.5; DB 2; Length 408;
Best Local Similarity 28.8%; Pred. No. 6;
Matches 32; Conservative 16; Mismatches 44; Indels 19; Gaps 5;

QY 24 CKTANGTAIPIG--GGSANVYVNLAPVNVVGNQL--VVDLSQIFCHNDYP-----E 71
Db 287 CSMTSTTAIPPEIPSNVNGVQMAPGVINGRNLGSGVNCYQLATLGPYKIALKFTFFD 346

QY 72 TITDVTTLQSG--SAYGGVLSNFSCTVKYSGSYPPPTTSETPRVYVNSRTD 121
Db 347 TGNIDVTIYDGPSTESPILG-----RYSNNMYEFTTSGSMTVMVTFKSD 391

RESULT 15
A59200
acid phosphatase (EC 3.1.3.2) purple 1, precursor [validated] - sweet potato
N:Alternate names: purple acid phosphatase (PAP)
C:Species: Ipomoea batatas (sweet potato)
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: A59200; C59200
R:Schenk, G.; Ge, Y.; Carrington, L.E.; Wynne, C.J.; Searle, I.R.; Carroll, B.J.; Hamilt-
Arch. Biochem. Biophys. 370, 183-189, 1999
A:Title: Binuclear metal centers in plant purple acid phosphatases: Fe-Mn in sweet potat-
A:Reference number: Z25293; MUID:99441212; PMID:10510276
A:Accession: A59200
A:Molecule type: mRNA
A:Residues: 1-473 <KUR>
A:Cross-references: UNIPROT:Q9SE00; GB:AF200825; NID:G6635440; PIDN:AAF19821.1; PID:G663
A:Accession: C59200
A:Molecule type: protein
A:Residues: 39-61 <KU2>
C:Comment: Metal content is iron and manganese, with small amounts of iron and copper.
C:Complex: homodimer
C:Function:
A:Description: catalyzes the hydrolysis of phosphate monoesters [validated, MUID:9944121
A:Note: active on a variety of activated and unactivated phosphate ester and anhydride s
C:Superfamily: kidney bean purple acid phosphatase; phosphatase core homology
C:Keywords: glycoprotein; homodimer; iron; manganese; metalloprotein; phosphoric monoest
F:39-473/Product: purple acid phosphatase #status experimental <MAT>
F:166-240/Domain: phosphatase core homology <PEC>
F:118,180,311,433/Binding site: carbonylate (Asn) (covalent) #status predicted
F:172,201,204,362/Binding site: iron (Asp, Asp, Tyr, His) #status predicted
F:201,238,323,360/Binding site: manganese (Asp, Asp, His, His) #status predicted
F:239,333/Active site: His #status predicted
F:382/Disulfide bonds: interchain #status predicted

Query Match 9.8%; Score 80.5; DB 1; Length 473;
Best Local Similarity 37.9%; Pred. No. 11;
Matches 22; Conservative 12; Mismatches 19; Indels 5; Gaps 2;

QY 76 YVTLQSGVLSNFSCTVKYSGSYPPPTTSETPRVYVNSRTDGPVALYLTVP 133
Db 274 WYAIKRASAHITVLSSYSGFVKYQKWF--TSELEKV---NRSETPWLVVHAPL 326

Search completed: November 27, 2004, 16:07:18
Job time : 43 secs

Mon Nov 29 14:12:21 2004

us-10-015-085-4_copy_1_158.rpr

Page 6

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: November 27, 2004, 16:07:26 ; Search time 141 Seconds

(without alignments)
397.459 Million cell updates/sec

Title: US-10-015-085-4_COPY_1_158

Perfect score: 821

Sequence: 1 MKRVITLPAVLLMGWSVNAW.....VAIKAGSLIAVILLRQTNNY 158

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

-Total number of hits satisfying chosen parameters: 1575965

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

-Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptcdat1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptcdat1/pubpaa/PTC_NEW_PUB.pep:*
- 3: /cgn2_6/ptcdat1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptcdat1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptcdat1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptcdat1/pubpaa/PTCUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptcdat1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptcdat1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptcdat1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptcdat1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptcdat1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptcdat1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptcdat1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptcdat1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptcdat1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptcdat1/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptcdat1/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptcdat1/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptcdat1/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptcdat1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	821	100.0	300	9	US-09-912-020-367
2	821	100.0	300	14	US-10-288-978-2
3	821	100.0	300	14	US-10-015-085-4
4	821	100.0	300	17	US-10-681-381B-1
5	809	98.5	300	15	US-10-607-834-14
6	712	86.7	268	10	US-09-747-802-1
7	712	86.7	268	16	US-10-789-619-1
8	712	86.7	279	9	US-09-900-575-34
9	712	86.7	279	9	US-09-900-575-44
10	709	86.4	279	9	US-09-900-575-42
11	708	86.2	279	9	US-09-900-575-27
12	708	86.2	279	9	US-09-900-575-36
13	708	86.2	279	9	US-09-900-575-39

14	708	86.2	279	9	US-09-900-575-55	Sequence 55, Appl
15	706	86.0	279	9	US-09-900-575-23	Sequence 23, Appl
16	705	85.9	279	9	US-09-900-575-29	Sequence 29, Appl
17	704	85.7	279	9	US-09-900-575-32	Sequence 32, Appl
18	704	85.7	279	9	US-09-900-575-37	Sequence 37, Appl
19	703	85.6	280	9	US-09-900-575-30	Sequence 30, Appl
20	702	85.5	279	9	US-09-900-575-43	Sequence 43, Appl
21	700	85.3	279	9	US-09-900-575-31	Sequence 31, Appl
22	700	85.3	279	9	US-09-900-575-38	Sequence 38, Appl
23	700	85.3	279	9	US-09-900-575-40	Sequence 40, Appl
24	700	85.3	279	15	US-10-607-834-15	Sequence 15, Appl
25	697	84.9	279	9	US-09-900-575-25	Sequence 25, Appl
26	697	84.9	279	9	US-09-900-575-26	Sequence 26, Appl
27	697	84.9	279	9	US-09-900-575-33	Sequence 33, Appl
28	697	84.9	279	9	US-09-900-575-45	Sequence 45, Appl
29	694	84.5	279	9	US-09-900-575-24	Sequence 24, Appl
30	694	84.5	279	9	US-09-900-575-28	Sequence 28, Appl
31	694	84.5	279	9	US-09-900-575-35	Sequence 35, Appl
32	690	84.0	279	9	US-09-900-575-41	Sequence 41, Appl
33	130	15.8	46	10	US-09-747-802-74	Sequence 74, Appl
34	130	15.8	46	16	US-10-789-619-74	Sequence 74, Appl
35	128	15.6	46	10	US-09-747-802-75	Sequence 75, Appl
36	128	15.6	46	16	US-10-789-619-75	Sequence 75, Appl
37	123	15.0	27	10	US-09-747-802-2	Sequence 2, Appl
38	123	15.0	27	16	US-10-789-619-2	Sequence 2, Appl
39	122	14.9	25	10	US-09-747-802-3	Sequence 3, Appl
40	122	14.9	25	16	US-10-789-619-3	Sequence 3, Appl
41	118	14.4	46	10	US-09-747-802-77	Sequence 77, Appl
42	118	14.4	46	16	US-10-789-619-77	Sequence 77, Appl
43	111	13.5	46	10	US-09-747-802-76	Sequence 76, Appl
44	111	13.5	46	16	US-10-789-619-76	Sequence 76, Appl
45	102	12.4	25	10	US-09-747-802-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-09-912-020-367
; Sequence 367, Application US/09912020
; Patent No. US20020045592A1
; GENERAL INFORMATION:
; APPLICANT: Zyskind, Judith
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Trawick, John
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Froelich, Jamie M.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; FILE REFERENCE: ESCHERICHIA COLI
; CURRENT APPLICATION NUMBER: US/09/912,020
; CURRENT FILING DATE: 2001-07-23
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/117,405
; PRIOR FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 367
; LENGTH: 300
; TYPE: PRT
; ORGANISM: E. Coli
US-09-912-020-367

Query Match 100.0%; Score 821; DB 9; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.2e-76;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLPAVLLMGWSVNAWVFACKTANGTAIP:GGGSANVYVNLAPVYVGNLWVLLS 60
|||||

Db 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVQNLVVDLS 60
QY 61 TOIFCHNDYPETITDYVTLQGSAYGGVLSNFSGTVKSGSYPPPTTSETPRVYNSRT 120
Db 61 TOIFCHNDYPETITDYVTLQGSAYGGVLSNFSGTVKSGSYPPPTTSETPRVYNSRT 120
QY 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNNY 158
Db 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 2

US-10-288-978-2
; Sequence 2, Application US/10288978
; Publication No. US20030099665A1
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Hultgren, Scott J.
; TITLE OF INVENTION: Chapterone and Adhesin Proteins; Vaccines, Diagnostics
; FILE OF INVENTION: Chapterone and Methods for Treating Infections
; FILE REFERENCE: 469201-362
; CURRENT APPLICATION NUMBER: US/10/288,978
; CURRENT FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: US/09/298,494
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: US 60/082,824
; PRIOR FILING DATE: 1998-04-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-288-978-2

Query Match 100.0%; Score 821; DB 14; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.2e-76;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVQNLVVDLS 60
Db 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVQNLVVDLS 60
QY 61 TOIFCHNDYPETITDYVTLQGSAYGGVLSNFSGTVKSGSYPPPTTSETPRVYNSRT 120
Db 61 TOIFCHNDYPETITDYVTLQGSAYGGVLSNFSGTVKSGSYPPPTTSETPRVYNSRT 120
QY 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNNY 158
Db 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 3

US-10-015-085-4
; Sequence 4, Application US/10015085
; Publication No. US20030199071A1
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon R.
; APPLICANT: Hultgren, Scott J.
; APPLICANT: Hung, Chia-Suei
; APPLICANT: Bouckaert, Julie
; TITLE OF INVENTION: Mutant Proteins, High Potency Inhibitory Antibodies, and FimCH
; FILE OF INVENTION: Crystall Structure
; FILE REFERENCE: 10271-037
; CURRENT APPLICATION NUMBER: US/10/015,085
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 300
; TYPE: PRT
; ORGANISM: E. coli

US-10-015-085-4

Query Match 100.0%; Score 821; DB 14; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.2e-76;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVQNLVVDLS 60
Db 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVQNLVVDLS 60
QY 61 TOIFCHNDYPETITDYVTLQGSAYGGVLSNFSGTVKSGSYPPPTTSETPRVYNSRT 120
Db 61 TOIFCHNDYPETITDYVTLQGSAYGGVLSNFSGTVKSGSYPPPTTSETPRVYNSRT 120
QY 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNNY 158
Db 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 4

US-10-681-381B-1
; Sequence 1, Application US/10681381B
; Publication No. US20040224400A1
; GENERAL INFORMATION:
; APPLICANT: Schembri, Mark A
; APPLICANT: Klemm, Per
; TITLE OF INVENTION: Novel multifunctional adhesin proteins and their display in
; FILE OF INVENTION: microbial cells
; FILE REFERENCE: 54259.000004
; CURRENT APPLICATION NUMBER: US/10/681,381B
; CURRENT FILING DATE: 2003-10-09
; PRIOR APPLICATION NUMBER: US 60/083,794
; PRIOR FILING DATE: 1998-05-01
; PRIOR APPLICATION NUMBER: DK 1998 00598
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: US 09/301,704
; PRIOR FILING DATE: 1999-04-29
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-681-381B-1

Query Match 100.0%; Score 821; DB 17; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.2e-76;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVQNLVVDLS 60
Db 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVQNLVVDLS 60
QY 61 TOIFCHNDYPETITDYVTLQGSAYGGVLSNFSGTVKSGSYPPPTTSETPRVYNSRT 120
Db 61 TOIFCHNDYPETITDYVTLQGSAYGGVLSNFSGTVKSGSYPPPTTSETPRVYNSRT 120
QY 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNNY 158
Db 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 5

US-10-607-834-14
; Sequence 14, Application US/10607834
; Publication No. US20040067544A1
; GENERAL INFORMATION:
; APPLICANT: Greenlee, Winner and Sullivan, PC
; APPLICANT: Vogel, Viola
; TITLE OF INVENTION: Use of Adhesion Molecules as Bond Stress-Enhanced Nanoscale Bindir
; FILE OF INVENTION: Switches
; FILE REFERENCE: 91-02
; CURRENT APPLICATION NUMBER: US/10/607,834

```
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: US 60/392,467
; PRIOR FILING DATE: 2002-06-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 31
; SEQ ID NO 14
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-607-834-14

Query Match      98.5%; Score 809; DB 15; Length 300;
Best Local Similarity 98.1%; Pred. No. 3.8e-75;
Matches 155; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLS 60
DB 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLS 60

QY 61 TQIFCHNDYPERITDVTIQRGSAYGVLNFSGTVKYSGSSYPPTTSETPRVYNSRT 120
DB 61 TQIFCHNDYPERITDVTIQRGSAYGVLNFSGTVKYSGSSYPPTTSETPRVYNSRT 120

QY 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNNY 158
DB 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 6
US-09-747-802-1
; Sequence 1, Application US/09747802
; Publication No. US20030027979A1
; GENERAL INFORMATION:
; APPLICANT: WANG, CHANG YI
; TITLE OF INVENTION: SYNTHETIC PEPTIDE COMPOSITION AS IMMUNOGENS FOR
; FILE REFERENCE: 1151-4165
; CURRENT APPLICATION NUMBER: US/09/747,802
; CURRENT FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-747-802-1

Query Match      86.7%; Score 712; DB 10; Length 268;
Best Local Similarity 100.0%; Pred. No. 3.6e-65;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLSQIFCHNDYPERITDVTIQR 81
DB 1 FACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLSQIFCHNDYPERITDVTIQR 60

QY 82 GSAYGVLNFSGTVKYSGSSYPPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 141
DB 61 GSAYGVLNFSGTVKYSGSSYPPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 120

QY 142 KAGSLIAVLILRQTNNY 158
DB 121 KAGSLIAVLILRQTNNY 137

RESULT 7
US-10-789-619-1
; Sequence 1, Application US/10789619
; Publication No. US20040141993A1
; GENERAL INFORMATION:
; APPLICANT: WANG, CHANG YI
; TITLE OF INVENTION: SYNTHETIC PEPTIDE COMPOSITION AS IMMUNOGENS FOR
; FILE REFERENCE: 1151-4165
```

```
; CURRENT APPLICATION NUMBER: US/10/789,619
; CURRENT FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-789-619-1

Query Match      86.7%; Score 712; DB 16; Length 268;
Best Local Similarity 100.0%; Pred. No. 3.6e-65;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLSQIFCHNDYPERITDVTIQR 81
DB 1 FACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLSQIFCHNDYPERITDVTIQR 60

QY 82 GSAYGVLNFSGTVKYSGSSYPPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 141
DB 61 GSAYGVLNFSGTVKYSGSSYPPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 120

QY 142 KAGSLIAVLILRQTNNY 158
DB 121 KAGSLIAVLILRQTNNY 137

RESULT 8
US-09-900-575-34
; Sequence 34, Application US/09900575
; Patent No. US20020150587A1
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-34

Query Match      86.7%; Score 712; DB 9; Length 279;
Best Local Similarity 100.0%; Pred. No. 3.8e-65;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLSQIFCHNDYPERITDVTIQR 81
DB 1 FACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLSQIFCHNDYPERITDVTIQR 60

QY 82 GSAYGVLNFSGTVKYSGSSYPPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 141
DB 61 GSAYGVLNFSGTVKYSGSSYPPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 120

QY 142 KAGSLIAVLILRQTNNY 158
DB 121 KAGSLIAVLILRQTNNY 137

RESULT 9
US-09-900-575-44
; Sequence 44, Application US/09900575
; Patent No. US20020150587A1
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
```

APPLICANT: Revel, Andrew
APPLICANT: Auguste, Christine
APPLICANT: Burlein, Jeanne
TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
FILE REFERENCE: 469201-549
CURRENT APPLICATION NUMBER: US/09/900,575
CURRENT FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: US/60/216,750
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PatentIn version 3.0
SEQ ID NO 44
LENGTH: 279
TYPE: PRT
ORGANISM: E. coli
US-09-900-575-44

Query Match 86.7%; Score 712; DB 9; Length 279;
Best Local Similarity 100.0%; Pred. No. 3.8e-65;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 22 FACTTANGTAIPIGGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDYVTIQR 81
DB 1 FACTTANGTAIPIGGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDYVTIQR 60
QY 82 GSAYGGVLSNFGSTVKYSGSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 141
DB 61 GSAYGGVLSNFGSTVKYSGSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNNY 158
DB 121 KAGSLIAVLILRQTNNY 137

RESULT 10
US-09-900-575-42
Sequence 42, Application US/09900575
Patent No. US20020150587A1
GENERAL INFORMATION:
APPLICANT: Langermann, Solomon
APPLICANT: Revel, Andrew
APPLICANT: Auguste, Christine
APPLICANT: Burlein, Jeanne
TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
FILE REFERENCE: 469201-549
CURRENT APPLICATION NUMBER: US/09/900,575
CURRENT FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: US/60/216,750
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PatentIn version 3.0
SEQ ID NO 42
LENGTH: 279
TYPE: PRT
ORGANISM: E. coli
US-09-900-575-42

Query Match 86.4%; Score 709; DB 9; Length 279;
Best Local Similarity 99.3%; Pred. No. 7.8e-65;
Matches 136; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 22 FACTTANGTAIPIGGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDYVTIQR 81
DB 1 FACTTANGTAIPIGGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDYVTIQR 60
QY 82 GSAYGGVLSNFGSTVKYSGSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 141
DB 61 GSAYGGVLSNFGSTVKYSGSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNNY 158
DB 121 KAGSLIAVLILRQTNNY 137

RESULT 11
US-09-900-575-27
Sequence 27, Application US/09900575
Patent No. US20020150587A1
GENERAL INFORMATION:
APPLICANT: Langermann, Solomon
APPLICANT: Revel, Andrew
APPLICANT: Auguste, Christine
APPLICANT: Burlein, Jeanne
TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
FILE REFERENCE: 469201-549
CURRENT APPLICATION NUMBER: US/09/900,575
CURRENT FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: US/60/216,750
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PatentIn version 3.0
SEQ ID NO 27
LENGTH: 279
TYPE: PRT
ORGANISM: E. coli
US-09-900-575-27

Query Match 86.2%; Score 708; DB 9; Length 279;
Best Local Similarity 99.3%; Pred. No. 9.9e-65;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 22 FACTTANGTAIPIGGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDYVTIQR 81
DB 1 FACTTANGTAIPIGGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDYVTIQR 60
QY 82 GSAYGGVLSNFGSTVKYSGSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 141
DB 61 GSAYGGVLSNFGSTVKYSGSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNNY 158
DB 121 KAGSLIAVLILRQTNNY 137

RESULT 12
US-09-900-575-36
Sequence 36, Application US/09900575
Patent No. US20020150587A1
GENERAL INFORMATION:
APPLICANT: Langermann, Solomon
APPLICANT: Revel, Andrew
APPLICANT: Auguste, Christine
APPLICANT: Burlein, Jeanne
TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
FILE REFERENCE: 469201-549
CURRENT APPLICATION NUMBER: US/09/900,575
CURRENT FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: US/60/216,750
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PatentIn version 3.0
SEQ ID NO 36
LENGTH: 279
TYPE: PRT
ORGANISM: E. coli
US-09-900-575-36

Query Match 86.2%; Score 708; DB 9; Length 279;
Best Local Similarity 99.3%; Pred. No. 9.9e-65;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 22 FACTTANGTAIPIGGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDYVTIQR 81
DB 1 FACTTANGTAIPIGGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDYVTIQR 60
QY 82 GSAYGGVLSNFGSTVKYSGSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 141

Db 61 GSAYGGVLSNFSGTVKYSYPPPTTSETPRVVYNSRTDKPWPVALYLTTPVSSAGGVAI 120
142 KAGSLIAVLILRQTNYY 158
Db 121 KAGSLIAVLILRQTNYY 137

RESULT 13
US-09-900-575-39
; Sequence 39, Application US/09900575
; Patent No. US20020150587A1
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-39

Query Match 86.2%; Score 708; DB 9; Length 279;
Best Local Similarity 99.3%; Pred. No. 9.9e-65;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 22 FACTTANGTAIPIGGGSANVYNLAPVNVGNLWDLSTQIFCHNDYPETITDYVTIQR 81
Db 1 FACTTANGTAIPIGGGSANVYNLAPVNVGNLWDLSTQIFCHNDYPETITDYVTIQR 60
QY 82 GSAYGGVLSNFSGTVKYSYPPPTTSETPRVVYNSRTDKPWPVALYLTTPVSSAGGVAI 141
Db 61 GSAYGGVLSNFSGTVKYSYPPPTTSETPRVVYNSRTDKPWPVALYLTTPVSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNYY 158
Db 121 KAGSLIAVLILRQTNYY 137

RESULT 14
US-09-900-575-55
; Sequence 55, Application US/09900575
; Patent No. US20020150587A1
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 55
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence of FimH proteins for SEQ ID NO: 23 to 45
US-09-900-575-55

Query Match 86.2%; Score 708; DB 9; Length 279;
Best Local Similarity 99.3%; Pred. No. 9.9e-65;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 22 FACTTANGTAIPIGGGSANVYNLAPVNVGNLWDLSTQIFCHNDYPETITDYVTIQR 81
Db 1 FACTTANGTAIPIGGGSANVYNLAPVNVGNLWDLSTQIFCHNDYPETITDYVTIQR 60
QY 82 GSAYGGVLSNFSGTVKYSYPPPTTSETPRVVYNSRTDKPWPVALYLTTPVSSAGGVAI 141
Db 61 GSAYGGVLSNFSGTVKYSYPPPTTSETPRVVYNSRTDKPWPVALYLTTPVSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNYY 158
Db 121 KAGSLIAVLILRQTNYY 137

RESULT 15
US-09-900-575-23
; Sequence 23, Application US/09900575
; Patent No. US20020150587A1
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-23

Query Match 86.0%; Score 706; DB 9; Length 279;
Best Local Similarity 99.3%; Pred. No. 1.6e-64;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 22 FACTTANGTAIPIGGGSANVYNLAPVNVGNLWDLSTQIFCHNDYPETITDYVTIQR 81
Db 1 FACTTANGTAIPIGGGSANVYNLAPVNVGNLWDLSTQIFCHNDYPETITDYVTIQR 60
QY 82 GSAYGGVLSNFSGTVKYSYPPPTTSETPRVVYNSRTDKPWPVALYLTTPVSSAGGVAI 141
Db 61 GSAYGGVLSNFSGTVKYSYPPPTTSETPRVVYNSRTDKPWPVALYLTTPVSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNYY 158
Db 121 KAGSLIAVLILRQTNYY 137

Search completed: November 27, 2004, 16:24:26
Job time : 142 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	821	100.0	300	4	US-09-298-494-2	Sequence 2, Appli
2	821	100.0	300	4	US-09-492-708A-367	Sequence 367, Appl
3	714.5	87.0	314	4	US-09-489-039A-11890	Sequence 11890, A
4	712	86.7	268	4	US-09-747-802-1	Sequence 1, Appli
5	712	86.7	279	4	US-09-900-575-34	Sequence 34, Appl
6	712	86.7	279	4	US-09-900-575-44	Sequence 44, Appl
7	709	86.4	279	4	US-09-900-575-42	Sequence 42, Appl
8	708	86.2	279	4	US-09-900-575-27	Sequence 27, Appl
9	708	86.2	279	4	US-09-900-575-36	Sequence 36, Appl
10	708	86.2	279	4	US-09-900-575-39	Sequence 39, Appl
11	708	86.2	279	4	US-09-900-575-55	Sequence 55, Appl
12	706	86.0	279	4	US-09-900-575-23	Sequence 23, Appl
13	705	85.9	279	4	US-09-900-575-29	Sequence 29, Appl
14	704	85.7	279	4	US-09-900-575-32	Sequence 32, Appl
15	704	85.7	279	4	US-09-900-575-37	Sequence 37, Appl
16	703	85.6	280	4	US-09-900-575-30	Sequence 30, Appl
17	702	85.5	279	4	US-09-900-575-43	Sequence 43, Appl
18	700	85.3	279	4	US-09-900-575-31	Sequence 31, Appl
19	700	85.3	279	4	US-09-900-575-38	Sequence 38, Appl
20	700	85.3	279	4	US-09-900-575-40	Sequence 40, Appl
21	697	84.9	279	4	US-09-900-575-25	Sequence 25, Appl
22	697	84.9	279	4	US-09-900-575-26	Sequence 26, Appl
23	697	84.9	279	4	US-09-900-575-33	Sequence 33, Appl
24	697	84.9	279	4	US-09-900-575-45	Sequence 45, Appl
25	694	84.5	279	4	US-09-900-575-24	Sequence 24, Appl
26	694	84.5	279	4	US-09-900-575-28	Sequence 28, Appl
27	694	84.5	279	4	US-09-900-575-35	Sequence 35, Appl

```
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; TITLE OF INVENTION: ESCHERICHIA COLI
; FILE REFERENCE: ELITRA.001A
; CURRENT APPLICATION NUMBER: US/09/492,709A
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 367
; LENGTH: 300
; TYPE: PRT
; ORGANISM: E. Coli
; US-09-492-709A-367

Query Match      100.0%; Score 821; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.1e-81;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MKRVITLPAVLGMGSMVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVDLS 60
Db      1 MKRVITLPAVLGMGSMVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVDLS 60

Qy      61 TQIFCHNDYPETITDVTYVTLQSGAYGGVLSNFSGTGKSGSYPPPTTSETPRVYNSR 120
Db      61 TQIFCHNDYPETITDVTYVTLQSGAYGGVLSNFSGTGKSGSYPPPTTSETPRVYNSR 120

Qy      121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILROTNYY 158
Db      121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILROTNYY 158

RESULT 3
US-09-489-039A-11890
; Sequence 11890, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11890
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-11890

Query Match      87.0%; Score 714.5; DB 4; Length 314;
Best Local Similarity 84.3%; Pred. No. 1.7e-69;
Matches 134; Conservative 12; Mismatches 12; Indels 1; Gaps 1;

Qy      1 MKRVITLPAVLGMGSMVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVDLS 59
Db      14 MKKIIPLFTLLGLMGSMVNAWSFACKTANGTAIPIGGGSANVYNLTPAVNVGNLVDLS 73

Qy      60 TQIFCHNDYPETITDVTYVTLQSGAYGGVLSNFSGTGKSGSYPPPTTSETPRVYNSR 119
Db      74 STQIFCHNDYPETITDVTYVTLQSGAYGGVLSNFSGTGKSGSYPPPTTSETPRVYNSR 133

Qy      120 TDKPWPVALYLPVSSAGGVAIKAGSLIAVLILROTNYY 158
Db      134 TDKPWPVALYLPVSSAGGVAITAGSLIAVLILROTNYY 172

RESULT 4
US-09-747-802-1
; Sequence 1, Application US/09747802
```

```
; Patent No. 6780969
; GENERAL INFORMATION:
; APPLICANT: WANG, CHANG YI
; TITLE OF INVENTION: SYNTHETIC PEPTIDE COMPOSITION AS IMMUNOGENS FOR
; TITLE OF INVENTION: PREVENTION OF URINARY TRACT INFECTION
; FILE REFERENCE: 1151-4165
; CURRENT APPLICATION NUMBER: US/09/747,802
; CURRENT FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-747-802-1

Query Match      86.7%; Score 712; DB 4; Length 268;
Best Local Similarity 100.0%; Pred. No. 2.5e-69;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      22 FACKTANGTAIPIGGGSANVYNLAPVNVGNLVDLS TQIFCHNDYPETITDVTYVTLOR 81
Db      1 FACKTANGTAIPIGGGSANVYNLAPVNVGNLVDLS TQIFCHNDYPETITDVTYVTLOR 60

Qy      82 GSAYGGVLSNFSGTGKSGSYPPPTTSETPRVYNSRDKPWPVALYLPVSSAGGVAI 141
Db      61 GSAYGGVLSNFSGTGKSGSYPPPTTSETPRVYNSRDKPWPVALYLPVSSAGGVAI 120

Qy      142 KAGSLIAVLILROTNYY 158
Db      121 KAGSLIAVLILROTNYY 137

RESULT 5
US-09-900-575-34
; Sequence 34, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
; US-09-900-575-34

Query Match      86.7%; Score 712; DB 4; Length 279;
Best Local Similarity 100.0%; Pred. No. 2.7e-69;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      22 FACKTANGTAIPIGGGSANVYNLAPVNVGNLVDLS TQIFCHNDYPETITDVTYVTLOR 81
Db      1 FACKTANGTAIPIGGGSANVYNLAPVNVGNLVDLS TQIFCHNDYPETITDVTYVTLOR 60

Qy      82 GSAYGGVLSNFSGTGKSGSYPPPTTSETPRVYNSRDKPWPVALYLPVSSAGGVAI 141
Db      61 GSAYGGVLSNFSGTGKSGSYPPPTTSETPRVYNSRDKPWPVALYLPVSSAGGVAI 120

Qy      142 KAGSLIAVLILROTNYY 158
Db      121 KAGSLIAVLILROTNYY 137
```

```
RESULT 6
US-09-900-575-44
; Sequence 44, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burllein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 44
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-44

Query Match      86.7%; Score 712; DB 4; Length 279;
Best Local Similarity 100.0%; Pred. No. 2.7e-69;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPICGGGSANVYNLAPVNVGNLVLDLSTQIFCHNDYPETITDYVTLQR 81
DB 1 FACKTANGTAIPICGGGSANVYNLAPVNVGNLVLDLSTQIFCHNDYPETITDYVTLQR 60

QY 82 GSAYGGVLSNFSGTVKYSGSSYPPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 141
DB 61 GSAYGGVLSNFSGTVKYSGSSYPPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 120

QY 142 KAGSLIAVLILRQTNNY 158
DB 121 KAGSLIAVLILRQTNNY 137

RESULT 7
US-09-900-575-42
; Sequence 42, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burllein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 42
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-42

Query Match      86.4%; Score 709; DB 4; Length 279;
Best Local Similarity 99.3%; Pred. No. 5.7e-69;
Matches 136; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPICGGGSANVYNLAPVNVGNLVLDLSTQIFCHNDYPETITDYVTLQR 81
DB 1 FACKTANGTAIPICGGGSANVYNLAPVNVGNLVLDLSTQIFCHNDYPETITDYVTLQR 60

QY 82 GSAYGGVLSNFSGTVKYSGSSYPPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 141
DB 61 GSAYGGVLSNFSGTVKYSGSSYPPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 120

QY 142 KAGSLIAVLILRQTNNY 158
DB 121 KAGSLIAVLILRQTNNY 137

RESULT 8
US-09-900-575-27
; Sequence 27, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burllein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 27
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-27

Query Match      86.2%; Score 708; DB 4; Length 279;
Best Local Similarity 99.3%; Pred. No. 7.3e-69;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPICGGGSANVYNLAPVNVGNLVLDLSTQIFCHNDYPETITDYVTLQR 81
DB 1 FACKTANGTAIPICGGGSANVYNLAPVNVGNLVLDLSTQIFCHNDYPETITDYVTLQR 60

QY 82 GSAYGGVLSNFSGTVKYSGSSYPPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 141
DB 61 GSAYGGVLSNFSGTVKYSGSSYPPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 120

QY 142 KAGSLIAVLILRQTNNY 158
DB 121 KAGSLIAVLILRQTNNY 137

RESULT 9
US-09-900-575-36
; Sequence 36, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burllein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 36
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-36

Query Match      86.2%; Score 708; DB 4; Length 279;
Best Local Similarity 99.3%; Pred. No. 7.3e-69;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 22 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVTLQR 81
Db 1 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVTLQR 60
QY 82 GSAYGGVLSNFGSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 141
Db 61 GSAYGGVLSNFGSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNYY 158
Db 121 KAGSLIAVLILRQTNYY 137

RESULT 10
US-09-900-575-39
; Sequence 39, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; PRIOR FILING DATE: 2001-07-06
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 39
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-39

Query Match 86.2%; Score 708; DB 4; Length 279;
Best Local Similarity 99.3%; Pred. No. 7.3e-69;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVTLQR 81
Db 1 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVTLQR 60
QY 82 GSAYGGVLSNFGSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 141
Db 61 GSAYGGVLSNFGSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNYY 158
Db 121 KAGSLIAVLILRQTNYY 137

RESULT 11
US-09-900-575-55
; Sequence 55, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; PRIOR FILING DATE: 2001-07-06
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 55
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-55

Query Match 86.2%; Score 708; DB 4; Length 279;
Best Local Similarity 99.3%; Pred. No. 7.3e-69;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVTLQR 81
Db 1 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVTLQR 60
QY 82 GSAYGGVLSNFGSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 141
Db 61 GSAYGGVLSNFGSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNYY 158
Db 121 KAGSLIAVLILRQTNYY 137

RESULT 12
US-09-900-575-23
; Sequence 23, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; PRIOR FILING DATE: 2001-07-06
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 23
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-23

Query Match 86.0%; Score 706; DB 4; Length 279;
Best Local Similarity 99.3%; Pred. No. 1.2e-68;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVTLQR 81
Db 1 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVTLQR 60
QY 82 GSAYGGVLSNFGSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 141
Db 61 GSAYGGVLSNFGSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNYY 158
Db 121 KAGSLIAVLILRQTNYY 137

RESULT 13
US-09-900-575-29
; Sequence 29, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; PRIOR FILING DATE: 2001-07-06
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 29
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-29

Query Match 86.2%; Score 708; DB 4; Length 279;
Best Local Similarity 99.3%; Pred. No. 7.3e-69;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 22 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVTLQR 81
Db 1 FACKTANGTAIPIGGSANVYNLAPVNVGNLVVDLSTQIFCHNDYPETITDVTLQR 60
QY 82 GSAYGGVLSNFGSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 141
Db 61 GSAYGGVLSNFGSGTVKYSYSSYPFPTTSETPRVYNSRTDKPWPVALYLTVPSSAGGVAI 120
QY 142 KAGSLIAVLILRQTNYY 158
Db 121 KAGSLIAVLILRQTNYY 137
```

```

; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-37

Query Match      85.7%; Score 704; DB 4; Length 279;
Best Local Similarity 98.5%; Pred. No. 2e-68;
Matches 135, Conservative 1; Mismatches 1; Indels 0; Gaps 0

Qy 22  PACTANGTAIPITGGGANYVYNLAPVNVNGQNLVLDLSTQIFCHNDYPETITDYVTLQR 81
      |||||
Db 1  PACTANGTAIPITGGGANYVYNLAPVNVNGQNLVLDLSTQIFCHNDYPETITDYVTLQR 60

Qy 82  GSAYGGVLSNFSGTVKYSGSSYPPTTSETPRVVYNSRTDKPWPVALYLTVPSSAGGVAI 141
      |||||
Db 61  GSAYGGVLSNFSGTVKEYSGSSYPPTTSETPRVVYNSRTDKPWPVALYLTVPSSAGGVAI 120

Qy 142 KAGSLIAVLILRQTNNY 158
Db 121 KAGSLIAVLILRQTNNY 137

```

Search completed: November 27, 2004, 16:14:07
Job time : 40 secs

```

; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/09/900, 575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-29

      85.9%;   Score 705;   DB 4;   Length 279;
Query Match
Best Local Similarity 98.5%;   Pred. No. 1.5e-68;
Matches 135;   Conservative 1;   Mismatches 1;   Indels 0;   Gaps 0

QY      22  FACKTANGTAIPGGGSANVYNLAPVNVGNQLVVDLSTQIFCHNDYPETITDVTTLQR 81
      |||
Db       1  FACKTANGTAIPGGGSANVYNLAPVNVGNQLVVDLSTQIFCHNDYPETITDVTTLQR 60
      |||

QY      82  GSAYGGLVSNFGSTVKYSGSSYPFTTSETPRVNVNSRTDKPWPVALYLTVPSSAGGVAI 141
      |||
Db       61  GSAYGGLVSNFGSTVKYSGSSYPFTTSETPRVNVNSRTDKPWPVALYLTVPSSAGGLVI 120
      |||

QY      142  KAGSLIAVLILRQTNNY 158
      |||
Db       121  KAGSLIAVLILRQTNNY 137
      |||

```

```

RESULT 14
US-09-900-575-32
; Sequence 32, Application US/09900575
; Patent No. 6737063
; GENERAL INFORMATION:
; APPLICANT: Langermann, Solomon
; APPLICANT: Revel, Andrew
; APPLICANT: Auguste, Christine
; APPLICANT: Burlein, Jeanne
; TITLE OF INVENTION: Film Adhesin Proteins and Methods of Use
; FILE REFERENCE: 469201-549
; CURRENT APPLICATION NUMBER: US/05/900,575
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/60/216,750
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32
; LENGTH: 279
; TYPE: PRT
; ORGANISM: E. coli
US-09-900-575-32

```

Query Match	85.7%;	Score 704;	DB 4;	Length 279;
Best Local Similarity	98.5%;	Prod. No. 2e-68;		
Matches 135;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	22	FACKTANGTAIPIGGGSANVYNLAPVNVNGQNLVVDLSTQIFCHNDYPEITIDVVTQLOR	81	
Db	1	FACKTANGTAIPIGGGSANVYNLAPVNVNGQNLVVDLSTQIFCHNDYPEITIDVVTQLOR	60	
Qy	82	GSAYGVLNFSFGTWKYSGSSYPFFTSTETPRVNVNSRTDKPWPVALYLTVPSSAGGVAI	141	
Db	61	GSAYGVLNFSFGTWKYSGSSYPFFTSTETPRVNVNSRTDKPWPVALYLTVPSSAGGVVI	120	
Qy	142	KAGSLIAVLILROTNNY	158	
Db	121	KAGSLIAVLILROTNNY	137	

RESULT 15
US-09-900-575-37
; Sequence 37, Application US/09900575

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein -2protein search, using sw model
Run on: November 27, 2004, 16:06:28 ; Search time 153 Seconds
(without alignments)
370.452 Million cell updates/sec

Title: US-10-015-085-4_COPY_1_158
Perfect score: 821
Sequence: 1 MKRVITLFAVLWGWSVNAW.....VAIKAGSLIAVLILRTNNY 158

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Genesecp_238ep04:*

- 1: Genesecp1980s:*
- 2: Genesecp1980s:*
- 3: Genesecp2000s:*
- 4: Genesecp2001s:*
- 5: Genesecp2002s:*
- 6: Genesecp2003as:*
- 7: Genesecp2003bs:*
- 8: Genesecp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	821	100.0	300	AAR76745	Aar76745 FimH prot
2	821	100.0	300	AAV59456	Aav59456 E. coli p
3	821	100.0	300	AAV59456	Aav59456 E. coli p
4	821	100.0	300	AAV59456	Aav59456 E. coli p
5	821	100.0	300	AAV59456	Aav59456 E. coli p
6	821	100.0	300	AAV59456	Aav59456 E. coli p
7	821	100.0	300	AAV59456	Aav59456 E. coli p
8	817	99.5	300	AAV59456	Aav59456 E. coli p
9	817	99.5	300	AAV59456	Aav59456 E. coli p
10	817	99.5	300	AAV59456	Aav59456 E. coli p
11	817	99.5	300	AAV59456	Aav59456 E. coli p
12	816	99.4	300	AAV59456	Aav59456 E. coli p
13	815	99.3	300	AAV59456	Aav59456 E. coli p
14	811	98.8	300	AAV59456	Aav59456 E. coli p
15	809	98.5	300	AAV59456	Aav59456 E. coli p
16	809	98.5	300	AAV59456	Aav59456 E. coli p
17	809	98.5	300	AAV59456	Aav59456 E. coli p
18	809	98.5	300	AAV59456	Aav59456 E. coli p
19	808	98.4	300	AAV59456	Aav59456 E. coli p
20	802	97.7	300	AAV59456	Aav59456 E. coli p
21	801	97.6	300	AAV59456	Aav59456 E. coli p
22	797	97.1	300	AAV59456	Aav59456 E. coli p
23	787	95.9	296	AAV59456	Aav59456 E. coli p
24	714.5	87.0	314	AB065373	Ab065373 Klebsiell
25	712	86.7	268	AB068154	Ab068154 Escherich

26	712	86.7	279	4	AAV72515	Aav72515 Escherich
27	712	86.7	279	4	AAV72515	Aav72515 Escherich
28	712	86.7	279	4	AAV72515	Aav72515 Escherich
29	712	86.7	279	4	AAV72515	Aav72515 Escherich
30	712	86.7	279	5	AAE18433	Aae18433 Escherich
31	712	86.7	296	4	AAV72515	Aav72515 Escherich
32	712	86.7	304	4	AAV72515	Aav72515 Escherich
33	709	86.4	279	5	AAE18435	Aae18435 Escherich
34	708	86.2	279	5	AAE18429	Aae18429 Escherich
35	708	86.2	279	5	AAE18417	Aae18417 Escherich
36	708	86.2	279	5	AAE18436	Aae18436 Escherich
37	708	86.2	279	5	AAE18426	Aae18426 Escherich
38	706	86.0	279	5	AAE18413	Aae18413 Escherich
39	705	85.9	279	5	AAE18419	Aae18419 Escherich
40	704	85.7	279	5	AAE18427	Aae18427 Escherich
41	704	85.7	279	5	AAE18422	Aae18422 Escherich
42	703	85.6	280	5	AAE18420	Aae18420 Escherich
43	702	85.5	279	5	AAE18432	Aae18432 Escherich
44	700	85.3	279	5	AAE18430	Aae18430 Escherich
45	700	85.3	279	5	AAE18428	Aae18428 Escherich

ALIGNMENTS

RESULT 1
AAR76745
ID AAR76745 standard; protein; 300 AA.
XX
AC AAR76745;
XX
DT 16-OCT-2003 (revised)
DT 13-MAR-1996 (first entry)
XX
DE FimH protein derived from E. coli K12 strain PC31.
XX
KW FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
KW FimA; FimF; FimG; receptor binding site.
XX
OS Escherichia coli K12; strain PC31.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT Protein /note= "Signal peptide"
FT /note= "Mature FimH"
XX
PN WO9520657-A1.
XX
PD 03-AUG-1995.
XX
PF 27-JAN-1995; 95WO-DK000042.
XX
PR 27-JAN-1994; 94US-00187166.
XX
PA (GXBI-) GX BIOSYSTEMS AS.
XX
PI Sokurenko EV, Hasty DL, Klemm P, Pallesen L, Molin S;
XX
DR WPI; 1995-275442/36.
XX
PT Receptor specific bacterial adhesins - useful for targeting active
XX compounds and microbial cells to locations of receptors.
XX
PS Example 1; Page 88-89; 152pp; English.
XX
CC This sequence represents the FimH protein from E. coli K12 strain PC31.
CC FimH is located at the tip of the type 1 fimbriae and also intercalated
CC at intervals in the fimbrial organelle. Most forms of the FimH adhesin
CC target, and bind to, oligosaccharide structures containing terminally
CC located alpha-D-mannoside residues. FimH contains 4 cysteine residues
CC assumed to direct folding of the molecule into distinct functional
CC domains. For comparison FimA and the minor components FimF and FimG only

CC have 2 cysteine residues. The localisation of the cysteine residues in
 CC FimH points to a tandem arrangement of two ancestral genes. Similar amino
 CC acids can be found in similar positions in the two halves of the FimH
 CC protein. The "midway" point is located roughly around residue 150 in the
 CC mature protein. The two halves or domains of FimH have evolved
 CC differently with the N-terminal section becoming the domain harbouring
 CC the receptor binding site, whereas the C-terminal sector became the
 CC domain of the molecule required for integration into the fimbrial
 CC organelle. This sequence and those given in AAR76763-76 may be used in
 CC the production of a variant FimH adhesin which may be useful for
 CC targeting active compounds and microbial cells to locations comprising
 CC selected receptors to which the adhesins bind. (Updated on 16-OCT-2003 to
 CC standardise OS field)

XX
 SQ Sequence 300 AA;

Query Match 100.0%; Score 821; DB 2; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2e-79;
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS 60
 DB 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS 60
 QY 61 TQIFCHNDYPETITDVTTLQGSAYGGVLSNFSFGTVKSGSYPPPTTSETPRVYNSRT 120
 DB 61 TQIFCHNDYPETITDVTTLQGSAYGGVLSNFSFGTVKSGSYPPPTTSETPRVYNSRT 120
 QY 121 DKPWPVALYITPVSSAGGVAIKAGSLIAVLILRQTNNY 158
 DB 121 DKPWPVALYITPVSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 2

AAAY59456
 ID AAY59456 standard; peptide; 300 AA.
 XX
 AC AAY59456;
 XX
 DT 29-MAR-2000 (first entry)
 XX
 DE E. coli PC31 FimH protein.
 XX
 KW Multifunctional adhesin protein; organic receptor; bioremediation;
 KW biosorption; organic pollutant; herbicide; pesticide; toxic compound;
 KW recycling; metal isolation; metal binding domain.
 XX
 OS Escherichia coli.
 XX
 PN WO9957276-A1.
 XX
 PD 11-NOV-1999.
 XX
 PF 21-APR-1999; 99MO-DX000223.
 XX
 PR 30-APR-1998; 98DK-00000598.
 PR 01-MAY-1998; 98US-0083794P.
 XX
 PA (GYRE-) GYRE LTD.
 XX
 PI Schembri MA, Klemm P;
 XX
 DR WPI; 2000-072233/06.
 XX
 DR Novel recombinant cells useful for bioremediation and recycling
 PT processes.
 XX
 PS Claim 5; Page 5; 60pp; English.

XX This sequence is the E. coli FimH protein, which is an adhesin protein.
 CC The invention relates to a recombinant cell expressing a multifunctional
 CC adhesin (MA) protein on its surface. The MA protein has at least one
 CC binding domain (BD1) capable of binding to an organic receptor, and at

CC least one binding domain (BD2) not naturally present in the adhesin, and
 CC can bind to a compound to which the naturally occurring adhesin protein
 CC does not substantially bind. Cells of the invention are used as
 CC bioremediation or biosorption means to separate undesired compounds such
 CC as organic pollutants including herbicides and pesticides, or toxic
 CC compounds such as heavy metals from the environment, or for isolating
 CC precious compounds such as metals for recycling purposes

XX
 SQ Sequence 300 AA;

Query Match 100.0%; Score 821; DB 3; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2e-79;
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS 60
 DB 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGQNLVVDLS 60
 QY 61 TQIFCHNDYPETITDVTTLQGSAYGGVLSNFSFGTVKSGSYPPPTTSETPRVYNSRT 120
 DB 61 TQIFCHNDYPETITDVTTLQGSAYGGVLSNFSFGTVKSGSYPPPTTSETPRVYNSRT 120
 QY 121 DKPWPVALYITPVSSAGGVAIKAGSLIAVLILRQTNNY 158
 DB 121 DKPWPVALYITPVSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 3

AAAB16009
 ID AAB16009 standard; protein; 300 AA.
 XX
 AC AAB16009;
 XX
 DT 05-OCT-2000 (first entry)
 XX
 DE E. coli proliferation associated protein sequence SEQ ID NO:367.
 XX
 KW Escherichia coli; E. coli; proliferation; inhibition; screening;
 KW antimicrobial; bacterial growth; antisense therapy; antibacterial.
 XX
 OS Escherichia coli.
 XX
 PN WO2000044906-A2.
 XX
 PD 03-AUG-2000.
 XX
 PF 27-JAN-2000; 2000MO-US002200.
 XX
 PR 27-JAN-1999; 99US-0117405P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Zyskind J, Ohlisen KL, Trawick J, Forsyth RA, Froelich JM;
 PI Carr GJ, Yamamoto RT, Xu HH;
 XX
 DR WPI; 2000-514822/46.
 DR N-PSDB; AAA66015.
 XX
 DR Novel polynucleotides and polypeptides associated with microorganism
 PT proliferation, used to identify inhibitors of bacterial growth and
 PT proliferation, for use in antisense therapy.
 XX
 PS Claim 11; Page 274-275; 316pp; English.

XX AAA65809 to AAA65889 and AAA66058 to AAA66138 represent nucleotide
 CC sequences derived from Escherichia coli which inhibit E. coli
 CC proliferation. AAA65890 to AAA66055 and AAA65886 to AAA66040 represent
 CC nucleotide and protein sequences associated with E. coli proliferation.
 CC AAA66056 and AAA66057 represent primers used for sequencing E. coli
 CC proliferation inhibiting nucleotide inserts in an example from the
 CC present invention. Methods from the present invention can be used to
 CC identify a proliferation- required gene in a microorganism, by contacting
 CC a microorganism with a proliferation-required gene activity inhibitory

XX Vaccine includes complex comprising FimC-FimH, or PapD-PapG in a carrier.
XX Example 12; Col 25-28; 19pp; English.
XX The invention relates to a vaccine comprising a complex comprising FimC-
XX FimH, or PapD-PapG, in a carrier. Also included is a method for
XX protecting against a bacterial infection caused by plus-bearing
XX bacteria, comprising administering the vaccine to a human at risk. The
XX vaccine is useful for eliciting production of an antibody against the
XX complex when administered to a patient. It is used for protecting against
XX bacterial infection, e.g. urinary tract infection (UTI) e.g. cystitis,
XX pyelonephritis, or a bladder infection. The present sequence represents
XX E. coli adhesin, FimH, used to make a vaccine of the invention
XX
XX Sequence 300 AA;
XX
XX Query Match 100.0%; Score 821; DB 6; Length 300;
XX Best Local Similarity 100.0%; Pred. No. 2e-79;
XX Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MKRVITLFAVLGMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLS 60
XX DB 1 MKRVITLFAVLGMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLS 60
XX
XX QY 61 TOIFCHNDYPETITDVTLQGSAYGGVLSNFSNFGTVKYSYGFPTTSEPRVYVNSRT 120
XX DB 61 TOIFCHNDYPETITDVTLQGSAYGGVLSNFSNFGTVKYSYGFPTTSEPRVYVNSRT 120
XX
XX QY 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158
XX DB 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158
XX
XX RESULT 6
XX ADC24381
XX ID ADC24381 standard; protein; 300 AA.
XX AC ADC24381;
XX XX
XX DT 18-DEC-2003 (first entry)
XX DE E. coli FimH protein.
XX KW FimH; bacterial chaperone; immunogenic fragment; adhesin protein;
XX vaccine; bacterial infection; FimH; mannose-binding fragment;
XX urinary tract infection; UTI; enterobacterial infection; FimC;
XX immunisation; adhesin/chaperone complex; bladder infection;
XX antibacterial; enzyme.
XX OS Escherichia coli.
XX FH Key Location/Qualifiers
XX FT Misc-difference 16 /note= "Encoded by tgc"
XX FT Misc-difference 151 /note= "Encoded by agg"
XX FT Misc-difference 208 /note= "Encoded by cgt"
XX FT Misc-difference 299 /note= "Encoded by tat"
XX
XX US2003099665-A1.
XX 29-MAY-2003.
XX 06-NOV-2002; 2002US-00288978.
XX 23-APR-1998; 98US-0082824P.
XX 23-APR-1999; 99US-00298494.
XX (LANG/) LANGERMANN S.
XX (HULT/) HULTGREN S J.
XX PA

PA (PINK/) PINKNER J S.
PA (AUGU/) AUGUSTE C G.
XX Langermann S, Hultgren SJ, Pinkner JS, Auguste CG;
XX MPI; 2003-678130/64.
XX N-PSDB; ADC24380.
XX New vaccine comprising bacterial chaperone protein or bacterial adhesin
XX protein FimH or a mannose-binding fragment of FimH, useful for preventing
XX or treating enterobacterial infections.
XX Example 12; SEQ ID NO 2; 21pp; English.
XX The invention discloses a vaccine against bacterial infections comprising
XX a complex of a bacterial chaperone protein with an adhesin protein or an
XX immunogenic fragment of the adhesin protein. Also claimed is a vaccine
XX against bacterial infections comprising the bacterial adhesin protein
XX FimH, or a mannose-binding fragment of FimH, an antibody raised against a
XX complex of a bacterial chaperone protein with either an adhesin or an
XX immunogenic fragment of the adhesin protein, an antibody raised against
XX either the bacterial adhesin protein FimH or an immunogenic mannose-
XX binding fragment of FimH, preventing or treating enterobacterial infections in a
XX (UTI) in a host, preventing or treating enterobacterial infections, e.g.
XX protein is FimC. Preventing or treating the host with the vaccine and at least
XX one antibody raised against a complex of a FimC with either FimH or an
XX immunogenic mannose-binding fragment of FimH. The protein complex is an
XX adhesin/chaperone complex. The vaccine is useful for preventing or
XX treating enterobacterial infections, particularly, the vaccine is used
XX for urinary tract or bladder infections caused by Escherichia coli. The
XX antibody raised against a complex of a bacterial chaperone protein is
XX useful for detecting urinary tract infections and for the prevention
XX and/or treatment of urinary tract infections caused by Escherichia coli.
XX The sequence presented is the E. coli FimH protein.
XX
XX Sequence 300 AA;
XX
XX Query Match 100.0%; Score 821; DB 7; Length 300;
XX Best Local Similarity 100.0%; Pred. No. 2e-79;
XX Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MKRVITLFAVLGMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLS 60
XX DB 1 MKRVITLFAVLGMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVQNLVVDLS 60
XX
XX QY 61 TOIFCHNDYPETITDVTLQGSAYGGVLSNFSNFGTVKYSYGFPTTSEPRVYVNSRT 120
XX DB 61 TOIFCHNDYPETITDVTLQGSAYGGVLSNFSNFGTVKYSYGFPTTSEPRVYVNSRT 120
XX
XX QY 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158
XX DB 121 DKPWPVALYLTVPSSAGGVAIKAGSLIAVLILRQTNYY 158
XX
XX RESULT 7
XX ADG17420
XX ID ADG17420 standard; protein; 300 AA.
XX AC ADG17420;
XX XX
XX DT 26-FEB-2004 (first entry)
XX DE Escherichia coli FimH D-mannose-binding adhesin wild-type protein.
XX KW urinary tract infection; UTI; FimH; antibacterial; virucide; bacterial;
XX viral infection; vaccine; FimH; D-mannose-binding adhesin; wild-type;
XX protein co-ordinate data.
XX OS Escherichia coli.
XX XX
XX WO2002102974-A2.
XX PN

```

XX PD 27-DEC-2002.
XX PF
XX PP
XX PR 10-DEC-2001; 2001WO-US047994.
XX PR 08-DEC-2000; 2000US-0254353P.
XX PR 29-JUN-2001; 2001US-0301878P.
XX PR
XX PA (MEDI-) MEDIMUNE INC.
XX PI
XX PI Langermann SR, Hultgren SJ, Hung C, Bouckaert J;
XX PI WPI; 2003-167503/16.
XX DR N-PSDB; ADG17419.
XX DR
XX PT Preventing, treating or ameliorating one or more symptoms of bacterial or
XX PT viral infection, particularly URI, using immunospecific antibodies that
XX PT bind to antigens of mutant FimH protein.
XX PS
XX PS Claim 53; SEQ ID NO 4; 1194pp; English.
XX CC
XX CC The invention relates to a novel method for preventing, treating or
XX CC ameliorating one or more symptoms associated with a urinary tract
XX CC infection (URI) in a human subject infected with Escherichia coli
XX CC comprising administering one or more antibodies that immunospecifically
XX CC bind to one or more antigens of a mutant FimH protein having one or more
XX CC amino acid substitutions. The method of the invention has antibacterial
XX CC and virucide applications and may be useful for the prevention, treatment
XX CC and/or amelioration of a bacterial or viral infection, in particular a
XX CC urinary tract infection, via the production of a vaccine. The current
XX CC sequence is that of the Escherichia coli FimH D-mannose-binding adhesin
XX CC wild-type protein of the invention.
XX SQ
XX SQ Sequence 300 AA;
XX
XX Query Match 100.0%; Score 821; DB 7; Length 300;
XX Best Local Similarity 100.0%; Pred. No. 2e-79;
XX Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
XX Db 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
XX
XX QY 61 TQIFCHNDYPETITDVTYTLQGSAYGGLVSNFSGTVKYGSSYPPTTSETPRVYNSRT 120
XX Db 61 TQIFCHNDYPETITDVTYTLQGSAYGGLVSNFSGTVKYGSSYPPTTSETPRVYNSRT 120
XX
XX QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILQTNYY 158
XX Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILQTNYY 158
XX
XX RESULT 8
XX ID AAR76774
XX AC AAR76774;
XX AC AAR76774;
XX DT 16-OCT-2003 (revised)
XX DT 15-MAR-1996 (first entry)
XX DE
XX DE FimH protein derived from E. coli clinical isolate MJH31-3.
XX KW FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
XX KW FimA; FimF; FimG; receptor binding site.
XX OS Escherichia coli; clinical isolate MJH31-3.
XX FH Key Location/Qualifiers
XX FT Peptide 1..21
XX FT /note= "Signal peptide"
XX FT Protein 22..300
XX FT /note= "Mature FimH"

```

```

XX WO9520657-A1.
XX 03-AUG-1995.
XX 27-JAN-1995; 9SWO-DK000042.
XX 27-JAN-1994; 94US-00187166.
XX (GXBI-) GX BIOSYSTEMS AS.
XX Sokurenko EV, Hasty DL, Klemm P, Pallesen L, Molin S;
XX WPI; 1995-275442/36.
XX N-PSDB; AAQ93068.
XX Receptor specific bacterial adhesins - useful for targeting active
XX compounds and microbial cells to locations of receptors.
XX Example 1; Page 44-45; 152pp; English.
XX
XX The sequences given in AAR76763-76 are FimH proteins from various E. coli
XX clinical isolates. FimH is located at the tip of the type 1 fimbriae and
XX also intercalated at intervals in the fimbrial organelle. Most forms of
XX the FimH adhesin target to, and bind to, oligosaccharide structures
XX containing terminally located alpha-D-mannoside residues. FimH contains 4
XX cysteine residues assumed to direct folding of the molecule into distinct
XX functional domains. For comparison FimA and the minor components FimF and
XX FimG only have 2 cysteine residues. The localisation of the cysteine
XX residues in FimH points to a tandem arrangement of two ancestral genes.
XX Similar amino acids can be found in similar positions in the two halves
XX of the FimH protein. The "midway" point is located roughly around residue
XX 150 in the mature protein. The two halves or domains of FimH have evolved
XX differently with the N-terminal section becoming the domain harbouring
XX the receptor binding site, whereas the C-terminal sector became the
XX domain of the molecule required for integration into the fimbrial
XX organelle. These sequences may be used in the production of a variant
XX FimH adhesin which may be useful for targeting active compounds and
XX microbial cells to locations comprising selected receptors to which the
XX adhesins bind. (Updated on 16-OCT-2003 to standardise OS field)
XX SQ
XX SQ Sequence 300 AA;
XX
XX Query Match 99.5%; Score 817; DB 2; Length 300;
XX Best Local Similarity 99.4%; Pred. No. 5.4e-79;
XX Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
XX Db 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYNLAPVNVGNLVVDLS 60
XX
XX QY 61 TQIFCHNDYPETITDVTYTLQGSAYGGLVSNFSGTVKYGSSYPPTTSETPRVYNSRT 120
XX Db 61 TQIFCHNDYPETITDVTYTLQGSAYGGLVSNFSGTVKYGSSYPPTTSETPRVYNSRT 120
XX
XX QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILQTNYY 158
XX Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILQTNYY 158
XX
XX RESULT 9
XX ID AAR76771
XX AC AAR76771 standard; protein; 300 AA.
XX AC AAR76771;
XX DT 16-OCT-2003 (revised)
XX DT 15-MAR-1996 (first entry)
XX DE
XX DE FimH protein derived from E. coli clinical isolate KS-54.
XX KW FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
XX KW FimA; FimF; FimG; receptor binding site.

```

OS Escherichia coli; clinical isolate KS-54.
 XX
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT /note= "Signal peptide"
 FT Protein 22..300
 FT /note= "Mature FimH"
 XX
 XX WO9520657-A1.
 FN
 XX
 XX 03-AUG-1995.
 PD
 XX 27-JAN-1995; 95WO-DK000042.
 RF
 XX 27-JAN-1994; 94US-00187166.
 PR
 XX (GXBI-) GX BIOSYSTEMS AS.
 PA
 XX Sokurenko EV, Hasty DL, Klemm P, Pallesen L, Molin S;
 PI
 XX WPI; 1995-275442/36.
 DR
 XX N-PSDB; AAQ93073.
 DR
 XX
 XX Receptor specific bacterial adhesins - useful for targetting active
 PT compounds and microbial cells to locations of receptors.
 PT
 XX
 PS Example 1; Page 44-45; 152pp; English.
 XX
 XX The sequences given in AAR76763-76 are FimH proteins from various E. coli
 CC clinical isolates. FimH is located at the tip of the type 1 fimbriae and
 CC also intercalated at intervals in the fimbrial organelle. Most forms of
 CC the FimH adhesin target to, and bind to, oligosaccharide structures
 CC containing terminally located alpha-D-mannoside residues. FimH contains 4
 CC cysteine residues assumed to direct folding of the molecule into distinct
 CC functional domains. For comparison FimA and the minor components FimF and
 CC FimG only have 2 cysteine residues. The localisation of the cysteine
 CC residues in FimH points to a tandem arrangement of two ancestral genes.
 CC Similar amino acids can be found in similar positions in the two halves
 CC of the FimH protein. The "midway" point is located roughly around residue
 CC 150 in the mature protein. The two halves or domains of FimH have evolved
 CC differently with the N-terminal section becoming the domain harbouring
 CC the receptor binding site, whereas the C-terminal sector became the
 CC domain of the molecule required for integration into the fimbrial
 CC organelle. These sequences may be used in the production of a variant
 CC FimH adhesin which may be useful for targetting active compounds and
 CC microbial cells to locations comprising selected receptors to which the
 CC adhesins bind. (Updated on 16-OCT-2003 to standardise OS field)
 XX
 XX Sequence 300 AA;
 Query Match 99.5%; Score 817; DB 2; Length 300;
 Best Local Similarity 99.4%; Pred. No. 5.4e-79;
 Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MKRVITLFAVLMLGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVGNQLWVDLS 60
 Db 1 MKRVITLFAVLMLGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVGNQLWVDLS 60
 QY 61 TQIFCHNDYPETITDYVTTLQGSAYGGVLSNFSGTVKSGSYPPFTTSETPRVYNSRT 120
 Db 61 TQIFCHNDYPETITDYVTTLQGSAYGGVLSNFSGTVKSGSYPPFTTSETPRVYNSRT 120
 QY 121 DKPWPVALYITPVSSAGGVAIKAGSLIAVLLRQTNNY 158
 Db 121 DKPWPVALYITPVSSAGGVAIKAGSLIAVLLRQTNNY 158
 RESULT 10
 AAR76763
 ID AAR76763 standard; protein; 300 AA.
 XX
 XX AAR76763;
 AC

XX 16-OCT-2003 (revised)
 DT 15-MAR-1996 (first entry)
 XX
 XX FimH protein derived from E. coli clinical isolate KB21.
 DE
 XX
 XX FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
 KW FimA; FimF; FimG; receptor binding site.
 KW
 XX Escherichia coli; clinical isolate KB21.
 OS
 XX
 XX Key Location/Qualifiers
 FT Peptide 1..21
 FT /note= "Signal peptide"
 FT Protein 22..300
 FT /note= "Mature FimH"
 XX
 XX WO9520657-A1.
 FN
 XX
 XX 03-AUG-1995.
 PD
 XX 27-JAN-1995; 95WO-DK000042.
 PF
 XX 27-JAN-1994; 94US-00187166.
 PR
 XX (GXBI-) GX BIOSYSTEMS AS.
 PA
 XX Sokurenko EV, Hasty DL, Klemm P, Pallesen L, Molin S;
 PI
 XX WPI; 1995-275442/36.
 DR
 XX Receptor specific bacterial adhesins - useful for targetting active
 PT compounds and microbial cells to locations of receptors.
 PT
 XX
 PS Example 1; Page 44-45; 152pp; English.
 XX
 XX The sequences given in AAR76763-76 are FimH proteins from various E. coli
 CC clinical isolates. FimH is located at the tip of the type 1 fimbriae and
 CC also intercalated at intervals in the fimbrial organelle. Most forms of
 CC the FimH adhesin target to, and bind to, oligosaccharide structures
 CC containing terminally located alpha-D-mannoside residues. FimH contains 4
 CC cysteine residues assumed to direct folding of the molecule into distinct
 CC functional domains. For comparison FimA and the minor components FimF and
 CC FimG only have 2 cysteine residues. The localisation of the cysteine
 CC residues in FimH points to a tandem arrangement of two ancestral genes.
 CC Similar amino acids can be found in similar positions in the two halves
 CC of the FimH protein. The "midway" point is located roughly around residue
 CC 150 in the mature protein. The two halves or domains of FimH have evolved
 CC differently with the N-terminal section becoming the domain harbouring
 CC the receptor binding site, whereas the C-terminal sector became the
 CC domain of the molecule required for integration into the fimbrial
 CC organelle. These sequences may be used in the production of a variant
 CC FimH adhesin which may be useful for targetting active compounds and
 CC microbial cells to locations comprising selected receptors to which the
 CC adhesins bind. (Updated on 16-OCT-2003 to standardise OS field)
 XX
 XX Sequence 300 AA;
 Query Match 99.5%; Score 817; DB 2; Length 300;
 Best Local Similarity 99.4%; Pred. No. 5.4e-79;
 Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MKRVITLFAVLMLGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVGNQLWVDLS 60
 Db 1 MKRVITLFAVLMLGWSVNAWSFACKTANGTAIPIGGSANVYVNLAPVNVGNQLWVDLS 60
 QY 61 TQIFCHNDYPETITDYVTTLQGSAYGGVLSNFSGTVKSGSYPPFTTSETPRVYNSRT 120
 Db 61 TQIFCHNDYPETITDYVTTLQGSAYGGVLSNFSGTVKSGSYPPFTTSETPRVYNSRT 120
 QY 121 DKPWPVALYITPVSSAGGVAIKAGSLIAVLLRQTNNY 158
 Db 121 DKPWPVALYITPVSSAGGVAIKAGSLIAVLLRQTNNY 158

```

RESULT 11
AAR76773
ID AAR76773 standard; protein; 300 AA.
XX AC AAR76773;
XX XX 16-OCT-2003 (revised)
XX DT 15-MAR-1996 (first entry)
XX XX
XX FimH protein derived from E. coli clinical isolate MJ#9-3.
XX XX FimH; type 1 fimbriae; organella; adhesin; alpha-D-mannoside residue;
XX KW FimA; FimF; FimG; receptor binding site.
XX XX
XX OS Escherichia coli; clinical isolate MJ#9-3.
XX XX
XX Key a Location/Qualifiers
XX Peptide 1..21
XX /note= "Signal peptide"
XX FT Protein 22..300
XX /note= "Mature FimH"
XX FT
XX XX
XX PN W09520657-A1.
XX XX
XX PD 03-AUG-1995.
XX XX
XX PF 27-JAN-1995; 95WO-DK000042.
XX XX
XX PR 27-JAN-1994; 94US-00187166.
XX XX
XX PA (GXBI-) GX BIOSYSTEMS AS.
XX XX
XX PI Sokurenko EV, Hasty DL, Klemm P, Pallesen L, Molin S;
XX XX WPI; 1995-275442/36.
XX DR N-PSDB; AAQ93067.
XX XX
XX PT Receptor specific bacterial adhesins - useful for targetting active
XX compounds and microbial cells to locations of receptors.
XX XX
XX PS Example 1; Page 44-45; 152pp; English.
XX XX
XX CC The sequences given in AAR76763-76 are FimH proteins from various E. coli
XX clinical isolates. FimH is located at the tip of the type 1 fimbriae and
XX also intercalated at intervals in the fimbrial organelle. Most forms of
XX the FimH adhesin target to, and bind to, oligosaccharide structures
XX containing terminally located alpha-D-mannoside residues. FimH contains 4
XX cysteine residues assumed to direct folding of the molecule into distinct
XX functional domains. For comparison FimA and the minor components FimF and
XX FimG only have 2 cysteine residues. The localisation of the cysteine
XX residues in FimH points to a tandem arrangement of two ancestral genes.
XX Similar amino acids can be found in similar positions in the two halves
XX of the FimH protein. The "midway" point is located roughly around residue
XX 150 in the mature protein. The two halves or domains of FimH have evolved
XX differently with the N-terminal section becoming the domain harbouring
XX the receptor binding site, whereas the C-terminal sector became the
XX domain of the molecule required for integration into the fimbrial
XX organelle. These sequences may be used in the production of a variant
XX FimH adhesin which may be useful for targetting active compounds and
XX microbial cells to locations comprising selected receptors to which the
XX adhesins bind. (Updated on 16-OCT-2003 to standardise OS field)
XX XX
XX SQ Sequence 300 AA;
XX
XX Query Match 99.5%; Score 817; DB 2; Length 300;
XX Best Local Similarity 99.4%; Pred. No. 5.4e-79;
XX Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 MKRVITLFAVLGMGWSVNAWSACKTANGTAIPIGGGSANVYNLAPVAVNGQNLVVDLS 60
XX DB 1 MKRVITLFAVLGMGWSVNAWSACKTANGTAIPIGGGSANVYNLAPVAVNGQNLVVDLS 60

```

```

QY 61 TOIFCHNDYPETITDYTLQSGSAYGGVLSNFSGTVKYSYGFPTTSETPRVVNSRT 120
DB 61 TOIFCHNDYPETITDYTLQSGSAYGGVLSNFSGTVKYSYGFPTTSETPRVVNSRT 120
QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
DB 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
RESULT 12
ADG17467
ID ADG17467 standard; protein; 300 AA.
XX AC ADG17467;
XX XX 26-FEB-2004 (first entry)
XX DT
XX DE Escherichia coli FimH D-mannose-binding adhesin mutant protein Q154N.
XX KW urinary tract infection; UTI; FimH; antibacterial; virucide; bacterial;
XX KW viral infection; vaccine; FimH; D-mannose-binding adhesin; mutant;
XX KW mutein; protein co-ordinate data.
XX OS Escherichia coli.
XX XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 154
XX /note= "Wild-type Gln substituted for Asn"
XX XX
XX PN W02002102974-A2.
XX XX
XX PD 27-DEC-2002.
XX XX
XX PF 10-DEC-2001; 2001WO-US047994.
XX XX
XX PR 08-DEC-2000; 2000US-0254353P.
XX PR 29-JUN-2001; 2001US-0301878P.
XX XX
XX PA (MEDI-) MEDIMMUNE INC.
XX XX
XX PI Langermann SR, Hultgren SJ, Hung C, Bouckaert J;
XX XX WPI; 2003-167503/16.
XX XX
XX PT Preventing, treating or ameliorating one or more symptoms of bacterial or
XX viral infection, particularly UTI, using immunospecific antibodies that
XX bind to antigens of mutant FimH protein.
XX XX
XX PS Claim 52; Page; 1194pp; English.
XX XX
XX CC The invention relates to a novel method for preventing, treating or
XX ameliorating one or more symptoms associated with a urinary tract
XX infection (UTI) in a human subject infected with Escherichia coli
XX comprising administering one or more antibodies that immunospecifically
XX bind to one or more antigens of a mutant FimH protein having one or more
XX amino acid substitutions. The method of the invention has one or more
XX and/or virucide applications and may be useful for the prevention, treatment
XX and/or amelioration of a bacterial or viral infection, in particular a
XX urinary tract infection, via the production of a vaccine. The current
XX sequence is that of the Escherichia coli FimH D-mannose-binding adhesin
XX mutant protein of the invention which contains a Q154N mutation. This
XX sequence is not shown within the specification per se but was created by
XX the indexer used information from SEQ ID 4.
XX XX
XX SQ Sequence 300 AA;
XX
XX Query Match 99.4%; Score 816; DB 7; Length 300;
XX Best Local Similarity 99.4%; Pred. No. 6.9e-79;
XX Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 MKRVITLFAVLGMGWSVNAWSACKTANGTAIPIGGGSANVYNLAPVAVNGQNLVVDLS 60
XX DB 1 MKRVITLFAVLGMGWSVNAWSACKTANGTAIPIGGGSANVYNLAPVAVNGQNLVVDLS 60

```

Db 1 MKRVITLFAVLGMGWSVNAWSFACKTANGTAIPIGGSSANVYNLAPVNVNQNLVVDLS 60
QY 61 TQIFCHNDYPETITDYVTLQSGSAYGGVLSNFSGTVKYSGSSYPPTTSETPRVYNSRT 120
Db 61 TQIFCHNDYPETITDYVTLQSGSAYGGVLSNFSGTVKYSGSSYPPTTSETPRVYNSRT 120
QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRTNNY 158
Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRTNNY 158

RESULT 13
AAR76768
ID AAR76768 standard; protein; 300 AA.
XX AC AAR76768;
XX 16-OCT-2003 (revised)
DT 15-MAR-1996 (first entry)
XX FimH protein derived from E. coli clinical isolate CSH-50.
XX FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
KW FimA; FimF; FimG; receptor binding site.
XX Escherichia coli; clinical isolate CSH-50.
XX Key Location/Qualifiers
FT Peptide 1..21 /note= "Signal peptide"
FT Protein 22..300
FT /note= "Mature FimH"
FN WO9520657-A1.
XX 03-AUG-1995.
XX 27-JAN-1995; 95WO-DK000042.
XX 27-JAN-1994; 94US-00187166.
XX (GXBI-) GX BIOSYSTEMS AS.
XX Sokurenko EV, Hasty DL, Klemm P, Pallesen L, Molin S;
XX WPI; 1995-275442/36.
XX Receptor specific bacterial adhesins - useful for targetting active
PT compounds and microbial cells to locations of receptors.
XX Example 1; Page 44-45; 152pp; English.
XX The sequences given in AAR76763-76 are FimH proteins from various E. coli
CC clinical isolates. FimH is located at the tip of the type 1 fimbriae and
CC also intercalated at intervals in the fimbrial organelle. Most forms of
CC the FimH adhesin target to, and bind to, oligosaccharide structures
CC containing terminally located alpha-D-mannoside residues. FimH contains 4
CC cysteine residues assumed to direct folding of the molecule into distinct
CC functional domains. For comparison FimA and the minor components FimF and
CC FimG only have 2 cysteine residues. The localisation of the cysteine
CC residues in FimH points to a tandem arrangement of two ancestral genes.
CC Similar amino acids can be found in similar positions in the two halves
CC of the FimH protein. The "midway" point is located roughly around residue
CC 150 in the mature protein. The two halves or domains of FimH have evolved
CC differently with the N-terminal section becoming the domain harbouring
CC the receptor binding site, whereas the C-terminal sector became the
CC domain of the molecule required for integration into the fimbrial
CC organelle. These sequences may be used in the production of a variant
CC FimH adhesin which may be useful for targetting active compounds and
CC microbial cells to locations comprising selected receptors to which the
CC adhesins bind. (Updated on 18-OCT-2003 to standardise OS field)
XX Sequence 300 AA;

Query Match 99.3%; Score 815; DB 2; Length 300;
Best Local Similarity 99.4%; Pred. No. 8.8e-79;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKRVITLFAVLGMGWSVNAWSFACKTANGTAIPIGGSSANVYNLAPVNVNQNLVVDLS 60
Db 1 MKRVITLFAVLGMGWSVNAWSFACKTANGTAIPIGGSSANVYNLAPVNVNQNLVVDLS 60
QY 61 TQIFCHNDYPETITDYVTLQSGSAYGGVLSNFSGTVKYSGSSYPPTTSETPRVYNSRT 120
Db 61 TQIFCHNDYPETITDYVTLQSGSAYGGVLSNFSGTVKYSGSSYPPTTSETPRVYNSRT 120
QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRTNNY 158
Db 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRTNNY 158

RESULT 14
AAR76772
ID AAR76772 standard; protein; 300 AA.
XX AC AAR76772;
XX 16-OCT-2003 (revised)
DT 15-MAR-1996 (first entry)
XX FimH protein derived from E. coli clinical isolate U221-3.
XX FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
KW FimA; FimF; FimG; receptor binding site.
XX Escherichia coli; clinical isolate U221-3.
XX Key Location/Qualifiers
FT Peptide 1..21 /note= "Signal peptide"
FT Protein 22..300 /note= "Mature FimH"
FN WO9520657-A1.
XX 03-AUG-1995.
XX 27-JAN-1995; 95WO-DK000042.
XX 27-JAN-1994; 94US-00187166.
XX (GXBI-) GX BIOSYSTEMS AS.
XX Sokurenko EV, Hasty DL, Klemm P, Pallesen L, Molin S;
XX WPI; 1995-275442/36.
XX N-PSDB; AAQ93074.
XX Receptor specific bacterial adhesins - useful for targetting active
PT compounds and microbial cells to locations of receptors.
XX Example 1; Page 44-45; 152pp; English.
XX The sequences given in AAR76763-76 are FimH proteins from various E. coli
CC clinical isolates. FimH is located at the tip of the type 1 fimbriae and
CC also intercalated at intervals in the fimbrial organelle. Most forms of
CC the FimH adhesin target to, and bind to, oligosaccharide structures
CC containing terminally located alpha-D-mannoside residues. FimH contains 4
CC cysteine residues assumed to direct folding of the molecule into distinct
CC functional domains. For comparison FimA and the minor components FimF and
CC FimG only have 2 cysteine residues. The localisation of the cysteine
CC residues in FimH points to a tandem arrangement of two ancestral genes.
CC Similar amino acids can be found in similar positions in the two halves
CC of the FimH protein. The "midway" point is located roughly around residue
CC 150 in the mature protein. The two halves or domains of FimH have evolved
CC differently with the N-terminal section becoming the domain harbouring

CC the receptor binding site, whereas the C-terminal sector became the
 CC domain of the molecule required for integration into the fimbrial
 CC organelle. These sequences may be used in the production of a variant
 CC FimH adhesin which may be useful for targeting active compounds and
 CC microbial cells to locations comprising selected receptors to which the
 CC adhesins bind. (Updated on 16-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 300 AA;

Query Match 98.8%; Score 811; DB 2; Length 300;
 Best Local Similarity 98.7%; Pred. No. 2.4e-78;
 Matches 155; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60
 DB 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60

QY 61 TQIFCHNDYPETITDYVTLQGSAYGGVLSNFGSTVKYSGSYPPFTTSETPRVYNSRT 120
 DB 61 TQIFCHNDYPETITDYVTLQGSAYGGVLSNFGSTVKYSGSYPPFTTSETPRVYNSRT 120

QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
 DB 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158

RESULT 15
 AAR76770
 ID AAR76770 standard; protein; 300 AA.
 XX AAR76770;
 AC
 XX
 DT 16-OCT-2003 (revised)
 DT 15-MAR-1996 (first entry)
 XX
 XX FimH protein derived from E. coli clinical isolate F-18.
 KW FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
 KW FimA; FimF; FimG; receptor binding site.
 XX
 OS Escherichia coli; clinical isolate F-18.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT Protein /note= "Signal peptide"
 FT Protein 22..300
 FT Protein /note= "Mature FimH"
 XX
 PN WO9520657-A1.
 XX
 XX 03-AUG-1995.
 XX
 XX 27-JAN-1995; 95WO-DK000042.
 XX
 XX 27-JAN-1994; 94US-00187166.
 XX
 XX (GXBI-) GX BIOSYSTEMS AS.
 XX
 XX Sokurenko EV, Hasty DL, Klemm P, Pallesen L, Molin S;
 XX
 XX WPI: 1995-275442/36.
 XX N-PSDB; AAQ93063.
 XX
 XX Receptor specific bacterial adhesins - useful for targeting active
 XX compounds and microbial cells to locations of receptors.
 XX
 XX Example_1; Page 44-45; 152pp; English.
 XX
 CC The sequences given in AAR7673-76 are FimH proteins from various E. coli
 CC clinical isolates. FimH is located at the tip of the type 1 fimbriae and
 CC also intercalated at intervals in the fimbrial organelle. Most forms of
 CC the FimH adhesin target to, and bind to, oligosaccharide structures
 CC containing terminally located alpha-D-mannoside residues. FimH contains 4

CC cysteine residues assumed to direct folding of the molecule into distinct
 CC functional domains. For comparison FimA and the minor components FimF and
 CC FimG only have 2 cysteine residues. The localisation of the cysteine
 CC residues in FimH points to a tandem arrangement of two ancestral genes.
 CC Similar amino acids can be found in similar positions in the two halves
 CC of the FimH protein. The "midway" point is located roughly around residue
 CC 150 in the mature protein. The two halves or domains of FimH have evolved
 CC differently with the N-terminal section becoming the domain harbouring
 CC the receptor binding site, whereas the C-terminal sector became the
 CC domain of the molecule required for integration into the fimbrial
 CC organelle. These sequences may be used in the production of a variant
 CC FimH adhesin which may be useful for targeting active compounds and
 CC microbial cells to locations comprising selected receptors to which the
 CC adhesins bind. (Updated on 16-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 300 AA;

Query Match 98.5%; Score 809; DB 2; Length 300;
 Best Local Similarity 98.1%; Pred. No. 3.9e-78;
 Matches 155; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60
 DB 1 MKRVITLFAVLLMGWSVNAWSFACKTANGTAIPIGGGSANVYVNLAPVNVGNLVVDLS 60

QY 61 TQIFCHNDYPETITDYVTLQGSAYGGVLSNFGSTVKYSGSYPPFTTSETPRVYNSRT 120
 DB 61 TQIFCHNDYPETITDYVTLQGSAYGGVLSNFGSTVKYSGSYPPFTTSETPRVYNSRT 120

QY 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158
 DB 121 DKPWPVALYLPVSSAGGVAIKAGSLIAVLILRQTNNY 158

Search completed: November 27, 2004, 16:13:23
 Job time : 155 secs

